The Linguistics of the Environment:

A Corpus-based Study of German Precipitation Constructions

by

Claudia Heinrich

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

University of Alberta

© Claudia Heinrich, 2024

Abstract

This dissertation provides a corpus-based constructional analysis of precipitation events in German (e.g., *es regnet* 'it rains') based on real-world data from *Sketch Engine*'s *deTenTen13* 16.5 billion-word web corpus. The research was sparked by a large, cross-linguistic survey of weather event encoding by Eriksen, Kittilä and Kolehmainen (2010, 2012, 2015), who propose a three-way typology of meteorological events (MEs) with a focus on precipitation. They suggest a typological scale of precipitation encoding, based on which element bears the semantic weight of marking the precipitate: the predicate, an argument, or a combination of both. The proposed types are **argument type** with a light verb (e.g., *Rain falls*), **predicate type** with a dummy subject (*It rains*), and **argument-predicate type** with a semantically heavy subject and predicate (*Rain rains/Snow rains*—the so-called cognate and split types, respectively). While they assert that languages universally possess the argument type of ME construction, they claim that Germanic languages typically adhere to the universally rare predicate type. While this may be true when speakers are commenting on immediately observable precipitation events, the corpus data show that German not only employs all kinds of precipitate encoding (e.g., *Regen fällt* 'rain is falling', *es regnet* 'it is raining', *Regen/Schnee regnet* 'rain/snow is raining'), but may add a second light argument in the argument constructions (e.g., *Es gibt Regen* 'it gives rain') or may even exploit other strategies by utilizing a prepositional phrase such as *bei Regen* 'while/if it's raining' (lit. 'at rain'). Additionally, there are culture-specific causal agents like *Frau Holle* 'Mother Hulda' and *Petrus* 'St. Peter', who are regularly employed as the entities responsible for certain types of precipitation.

My analysis demonstrates that the choice of ME construction depends on extra-propositional contextual factors such as tense/aspect marking, temporal and locative modification, quality and quantity specification, figurativity, presence of stance-taking elements,

genre and regional differences, as well as the individual precipitate type (i.e. rain, snow, sleet, hail). For example, the prevalent German structure *es gibt* 'there is' (lit. 'it gives') is known to express absolute existence (*Es gibt einen Gott* 'There is a God'). However, in the examined data, it is predominantly used in past and future tense (60%) while less than 20% of *es gibt Regen* 'there is rain' (lit. 'it gives rain') utterances express existential or generic propositions. At the same time, the PRECIPITATE+*fallen* ('PRECIPITATE +fall') argument-type construction seems to be chiefly weather reporting, as part of a technical meteorological genre, whereby information about the average weather in a specific region is conveyed, or in texts on "how weather works", which goes hand-in-hand with their predominant use in existential or generic propositions. This construction also appears frequently in literary genres. However, the literary genre also attracts the argument-predicate type of ME construction, as in the cognate type (e.g. *Grauer Regen regnet dir aufs Dach* 'Grey rain is raining onto your roof') or split type (e.g. *Herbstspaziergang bei dem es Schnee regnete* 'Autumn walk where it rained snow').

Ultimately, there is no homogeneous distribution characterizing the interaction of precipitation type and construction type. Correspondingly, we find that hail, for example, overwhelmingly appears in the "default" predicate construction when the event is figurative (*es hagelte Beschwerden* 'it hailed complaints') and largely imbued with negative semantic prosody. MEs expressing a literal use of hail prefer the argument construction. Consequently, the present investigation provides evidence that Erikson et al.'s (2010) typology is too coarse-grained with regard to German, and therefore needs refinement. In summary, this study emphasizes the diversity of constructions employed within the German expressive field of MEs, while providing a qualitative and quantitative analysis of how syntactic or constructional variation is driven by interaction with a number of lexical, extra-propositional, and extra-linguistic variables.

Acknowledgements

I would like to thank the following people, without whom I would not have been able to complete this research. First and foremost, I wish to express my gratitude to my “two mothers”: my *Doktormutter* Dr. Sally Rice for her enduring support, encouragement and patience during these long and challenging years and my birth mother, Urda Heinrich, for her love and unwavering support. I never would have got this far without you. I would also like to extend my deepest gratitude to my supervisory committee, Dr. Antti Arppe and Dr. Xiaoting Li, as well as my examiners, Dr. Laurent Gautier and Dr. Andrea Wilhelm, for their guidance, their helpful and thought-provoking comments, and the time and effort that they invested into the review of this manuscript. Last but not least, I would like to acknowledge the emotional support provided by my family and friends.

Table of Contents

Abstract ii

Acknowledgements iv

Table of Contents v

List of Tables xi

List of Figures xvii

1 Stormy Weather Expressions and Constructions 1

1.1 Introduction 1

1.2 Intransitives and Talmy's typology of motion events 4

1.3 Prior analyses of weather constructions 9

1.4 Below the intransitive: What is weathering? 30

1.5 Fundamentals of cognitive linguistic and usage-based approaches to the study of meteorological expressions and why I chose a corpus-based approach 34

1.6 The emerging field of lexical typology 46

1.7 The structure of this dissertation 53

2 Methodology and Data 55

2.1 Some general remarks on corpus linguistics 55

2.1.1 Two key concepts of corpus design: balancedness and representativeness 57

2.1.2 Making sense of corpus data 58

2.2 Data collection 61

2.2.1 The corpus of the German Web 2013 (*deTenTen13*) 63

2.2.2 Searching the corpus using CQL (Corpus Query Language) 68

2.2.3 Extracting the examples and randomizing the data 75

2.2.4 Constructing data frames for the extracted MEs and the coding scheme 78

2.2.5 The *DWDS* dictionary and corpora 84

2.2.6 The *DWDS* corpus query search 88

2.3 Preview of the analyses carried out in Chapter 3 and Chapter 4 91

3 Quantitative Analysis of the Target Constructions by Precipitate 94

3.1 Constructions with *Regen* 'rain' 95

3.1.1 The Argument Construction [*Regen+fallen*] 'rain+fall' 95

3.1.1.1 USAGE and GENRE variables 95

3.1.1.2 The TAM variable 98

3.1.1.3 The SPATIAL modifier variable 100

3.1.1.4 The NON-SPATIAL modifier variable 103

3.1.1.5 The co-ME variable 113

3.1.2 The Argument Construction [*Regen+geben*] 'rain+give' 114

3.1.2.1 USAGE and GENRE variables 114

3.1.2.2 The TAM variable 116

3.1.2.3 The SPATIAL modifier variable 118

3.1.2.4 The NON-SPATIAL modifier variable 120

3.1.2.5 The co-ME variable 127

3.1.3 The Predicate Construction [*es+regnen*] 'it+rain' 128

3.1.3.1 USAGE and GENRE variables 129

3.1.3.2 The TAM variable 130

3.1.3.3 The SPATIAL modifier variable 133

3.1.3.4 The NON-SPATIAL modifier variable 136

3.1.3.5 The co-ME variable 145

3.1.4 Distribution and frequencies of variables across the three constructions with the precipitate *Regen* 'rain' 146

3.1.4.1 USAGE and GENRE variables across the three constructions 147

3.1.4.2 The TAM variable across the three constructions 150

3.1.4.3 The SPATIAL modifier variable across the three constructions 154

3.1.4.4 The NON-SPATIAL modifier variable across the three constructions 157

3.1.5 Summary of *Regen* 'rain' constructions 160

3.2 Constructions with the precipitate *Schnee* 'snow' 163

3.2.1 The Argument Construction [*Schnee+fallen*] 'snow+fall' 164

3.2.1.1 USAGE and GENRE variables 165

3.2.1.2 The TAM variable 166

3.2.1.3 The SPATIAL modifier variable 167
3.2.1.4 The NON-SPATIAL modifier variable 169
3.2.2 The Argument Construction [*Schnee+geben*] 'snow+give' 172
3.2.2.1 USAGE and GENRE variables 173
3.2.2.2 The TAM variable 174
3.2.2.3 The SPATIAL modifier variable 175
3.2.2.4 The NON-SPATIAL modifier variable 176
3.2.3 The Predicate Construction [*es+schneien*] 'it+snow' 180
3.2.3.1 USAGE and GENRE variables 181
3.2.3.2 The TAM variable 183
3.2.3.3 The SPATIAL modifier variable 185
3.2.3.4 The NON-SPATIAL modifier variable 186
3.2.4 Distribution and frequencies of variables across the three constructions with the precipitate *Schnee* 'snow' 192
3.2.4.1 USAGE and GENRE variables across the three constructions 192
3.2.4.2 The TAM variable across the three constructions 195
3.2.4.3 The SPATIAL modifier variable across the three constructions 198
3.2.4.4 The NON-SPATIAL modifier variable across the three constructions 200
3.2.5 Summary of *Schnee* ‘snow’ constructions 203
3.3 Constructions with the precipitate *Schneeregen* 'sleet' 205
3.3.1 The Argument Construction [*Schneeregen+fallen*] 'sleet+fall' 206
3.3.1.1 USAGE and GENRE variables 206
3.3.1.2 The TAM variable 207
3.3.1.3 The SPATIAL modifier variable 209
3.3.1.4 The NON-SPATIAL modifier variable 211
3.3.2 The Argument Construction [*Schneeregen+geben*] 'sleet+give' 216
3.3.2.1 USAGE and GENRE variables 217
3.3.2.2 The TAM variable 218
3.3.2.3 The SPATIAL modifier variable 219
3.3.2.4 The NON-SPATIAL modifier variable 221
3.3.2.5 The co-ME variable 223

3.3.3 The Predicate Construction [*es+schneeregnen*] 'it+sleet' 225
3.3.3.1 USAGE and GENRE variables 226
3.3.3.2 The TAM variable 226
3.3.3.3 The SPATIAL modifier variable 227
3.3.3.4 The NON-SPATIAL modifier variable 228
3.3.4 Distribution and frequencies of variables across the three constructions with the precipitate *Schneeregen* 'sleet' 230
3.3.4.1 USAGE and GENRE variables across the three constructions 231
3.3.4.2 The TAM variable across the three constructions 232
3.3.4.3 The SPATIAL modifier variable across the three constructions 234
3.3.4.4 The NON-SPATIAL modifier variable across the three constructions 236
3.3.5 Summary of *Schneeregen* 'sleet' constructions 239
3.4 Constructions with the precipitate *Hagel* 'hail' 241
3.4.1 The Argument Construction [*Hagel+fallen*] 'hail+fall' 242
3.4.1.1 USAGE and GENRE variables 242
3.4.1.2 The TAM variable 244
3.4.1.3 The SPATIAL modifier variable 245
3.4.1.4 The NON-SPATIAL modifier variable 247
3.4.1.5 The co-ME variable: *mit Hagel* 'with hail' and *als Hagel* 'as hail' 252
3.4.2 The Argument Construction [*Hagel+geben*] 'hail+give' 253
3.4.2.1 USAGE and GENRE variables 253
3.4.2.2 The TAM variable 254
3.4.2.3 The SPATIAL modifier variable 256
3.4.2.4 The NON-SPATIAL modifier variable 258
3.4.2.5 The co-ME variable: *mit Hagel* 'with hail' and *als Hagel* 'as hail' 261
3.4.3 The Predicate Construction [*es+hagelt*] 'it+hail' 262
3.4.3.1 USAGE and GENRE variables 263
3.4.3.2 The TAM variable 264
3.4.3.3 The SPATIAL modifier variable 265
3.4.3.4 The NON-SPATIAL modifier variable 266

3.4.4 Distribution and frequencies of variables across the three constructions with the precipitate *Hagel* 'hail' 269
3.4.4.1 USAGE and GENRE variables across the three constructions 269
3.4.4.2 The TAM variable across the three constructions 272
3.4.4.3 The SPATIAL modifier variable across the three constructions 274
3.4.4.4 The NON-SPATIAL modifier variable across the three constructions 276
3.4.5 Summary of *Hagel* 'hail' constructions 280
3.5 Final summary of all three constructions with all four precipitates 283

4 Marked Weather Constructions 295
4.1 Introduction 295
4.2 Marked Argument-Predicate Constructions 297
4.2.1 The cognate type 298
4.2.2 The split type 312
4.2.2.1 Frequencies and usage of the precipitate verb *rieseln* 'flutter/trickle/drift' 312
4.2.2.2 Subjects occurring with the verb *rieseln* in the *deTenTen13* and *DWDS* corpora 317
4.3 Other marked constructions 329
4.3.1 Transitives with causal agents 329
4.3.1.1 *Frau Holle* 'Mother Hulda' 331
4.3.1.2 *Der heilige Petrus* 'Patron Saint Peter' 334
4.3.2 Prepositional constructions 338
4.3.3 Idiosyncratic MEs and -*wetter* 'weather' compounds 354
4.3.3.1 *Fritz-Walter-Wetter* and other "bad weather" compounds 355
4.3.3.2 *Kaiserwetter* and other "Bluebird Days" expressions 363
4.4 Figurative uses of conventional of MEs 367
4.5 Summary and conclusion 372

5 Whither the weather? 377
5.1 How we *actually* talk about the weather 378

5.2 Weather expressions in German beyond precipitation 383
5.3 Revisiting lexical typology and implications for *DaF* 'German as a foreign language' 386
5.4 Limitations of typologies and the present study 392
5.5 Implications for the lexicon-syntax interface 395
5.6 Summary and conclusion 396

References 399

List of Tables

1	Schematic illustration of the typology of meteorological expressions (adapted from Eriksen et al. 2010: Figure 1; [summaries mine])	15
2	Classification and frequency of temporal items in texts prepared for an expert audience vs. a lay audience (adapted from Liégeois 2021: 226.f.)	23
3	Total number of collocate co-occurrences with the verb lemma *regnen* 'to rain' (within a specified range of three words to the left and three words to the right of the lemma)	60
4	List of search and return features available in *Sketch Engine*	63
5	Composition counts of the German Web 2013 (deTenTen13) corpus	65
6	Makeup of the *deTenTen13* subcorpora	66
7	Examples of the [es geben] construction with the precipitate Regen 'rain' allowing between 0 to 6 words in between the target items	72
8	Schematic CQL searches carried out for precipitate type	75
9	Variables and variable levels for the annotation of the downloaded ME constructions analyzed in Chapters 3 and 4	80
10	Available DWDS corpora	86
11	Frequencies of literal and figurative USE for the [ARG+fallen-Cx construction and the precipitate Regen 'rain'	97
12	Raw and relative frequencies of the GENRE variable across the ARG+fallen-Cx and the precipitate Regen 'rain'	98
13	Raw and relative frequencies of the TAM variable across the ARG+fallen-Cx and the precipitate Regen 'rain'	100
14	Raw and relative frequencies of the SPATIAL modifier variable in the ARG+fallen construction and the precipitate Regen 'rain'	101
15	Raw and relative frequencies of NON-SPATIAL modification across ARG+fallen and the precipitate Regen 'rain'	104
16	Raw and relative frequencies of the co-ME variable across the ARG+fallen-Cx and the precipitate Regen 'rain'	113
17	Distribution and frequencies of the GENRE variable across the ARG+geben-Cx and the precipitate Regen 'rain'	115
18	Raw and relative frequencies of the TAM variable across the ARG+geben-Cx with the precipitate Regen 'rain'	116
19	Raw and relative frequencies for the SPATIAL modification variable across the ARG+geben-Cx and the precipitate Regen 'rain'	118
20	Raw and relative frequencies of the NON-SPATIAL modifier variable across the ARG+geben-Cx and the precipitate Regen 'rain'	120

21	Categories and frequencies of TEMPORAL modifiers occurring with the returns for the Regen+geben construction	122
22	Frequencies of modifier types and tokens in the Regen+geben returns specifying the quantity of rain as categorized by magnitude or absence of rainfall	123
23	Raw and relative frequencies of the co-ME variable across the ARG+geben-Cx and the precipitate Regen 'rain'	127
24	Frequencies of literal and figurative usage for the es+PRED-Cx and the precipitate Regen 'rain'	129
25	Distribution and frequencies of the GENRE variable across the es+PRED-Cx and the precipitate Regen 'rain'	129
26	Raw and relative frequencies of the TAM variable within the es+PRED-Cx and the precipitate Regen 'rain'	131
27	Distribution and frequencies of quasi-auxiliaries within returns of the es+PRED-Cx and the precipitate Regen 'rain'	131
28	Distribution and frequencies of spatial modifiers across the es+PRED-Cx with the precipitate Regen 'rain'	134
29	Raw and relative frequencies of the NON-SPATIAL modifier variable across the es+PRED-Cx and precipitate Regen 'rain'	137
30	Modifiers of QUANTITY across the es+PRED-Cx with the precipitate Regen 'rain'	140
31	Frequencies of the STANCE modifier and positive and negative SEMANTIC PROSODY across the es+PRED-Cx and the precipitate Regen 'rain'	144
32	Distribution and frequencies of the co-ME variable across the literal es+PRED-Cx and the precpitate Regen 'rain'	146
33	Frequencies of literal and figurative USE by construction type (ARG+fallen, ARG+geben, and es+PRED) with the precipitate Regen 'rain'	147
34	Raw and relative frequencies of the GENRE variable by construction type (ARG+fallen, ARG+geben, es+PRED) and the precipitate Regen 'rain'	148
35	Raw and relative frequencies of the TAM variable by construction type (ARG+fallen, ARG+geben, and es+PRED) and precipitate Regen 'rain' (fig/other removed)	151
36	Frequency and distribution of PAST tense differences across the three constructions (ARG+fallen, ARG+geben, and es+PRED) and the precipitate Regen 'rain'	152
37	Distribution and frequencies of the SPATIAL modification variable across constructions (ARG+fallen, ARG+geben, es+PRED) with the precipitate Regen 'rain'	155

38	Instances and frequencies of NON-SPATIAL modifier types across the three constructions (ARG+fallen, ARG+geben, es+PRED) with the precipitate Regen 'rain'	157
39	Frequencies of literal and figurative USAGE for the ARG+fallen-Cx and precipitate Schnee 'snow'	165
40	Raw and relative frequencies for the GENRE variable across the ARG+fallen-Cx and precipitate Schnee 'snow'	165
41	Raw and relative frequencies of the TAM variable across the ARG+fallen-Cx and precipitate Schnee 'snow'	166
42	Raw and relative frequencies of the SPATIAL modifier variable across ARG+fallen and the precipitate Schnee 'snow'	168
43	Raw and relative frequencies of the NON-SPATIAL modifier variable across the ARG+fallen-Cxx and the precipitate Schnee 'snow'	169
44	Frequencies of literal and figurative USE for the ARG+geben-Cx with the precipitate Schnee 'snow'	173
45	Raw and relative frequencies of the GENRE variable across the ARG+geben-Cx with the precipitate Schnee 'snow'	174
46	Raw and relative frequencies of the TAM variable for the ARG+geben-Cx with the precipitate Schnee 'snow'	175
47	Raw and relative frequencies of the SPATIAL modifier variable for the ARG+geben-Cx with the precipitate Schnee 'snow'	175
48	Raw and relative frequencies of the NON-SPATIAL modifier variable within the ARG+geben-Cx and the precipiate Schnee 'snow'	176
49	Frequencies of literal and figurative use for the es+PRED-Cx and the precipitate Schnee 'snow'	181
50	Raw and relative frequencies for the genre variable across the es+PRED-Cx and the precipitate Schnee 'snow'	182
51	Raw and relative frequencies of the TAM variable within the es+PRED-Cx and the precipitate Schnee 'snow' (non-literal instances removed)	183
52	Variations of tense and aspect within the data comprehensively coded as PAST for the es+PRED-Cx and the precipitate Schnee 'snow'	184
53	Raw and relative frequencies of the es+PRED-Cx for the SPATIAL modifier variable and the precipitate Schnee 'snow'	185
54	Raw and relative frequencies of the NON-SPATIAL MODIFIER variable (TEMP, QUANT, QUAL, STANCE) within the es+PRED-Cx and the precipitate Schnee 'snow' (literal use only)	187
55	Aggregate frequencies of literal and figurative USE by construction type (ARG+fallen, ARG+geben, and es+PRED) with the precipitate Schnee 'snow'	192
56	Raw and relative frequencies of the GENRE variable across the three constructions (ARG+fallen, ARG+geben, es+PRED) in combination with the precipitate Schnee 'snow'	193
57	Raw and relative frequencies of the TAM variable by construction type (ARG+fallen, ARG+geben, and es+PRED) with the precipitate Schnee 'snow' (FIGURATIVE/OTHER removed)	195

58	Overall distribution and absolute frequencies of the SPATIAL modifier variable within each individual construction (ARG+fallen, ARG+geben, and es+PRED) with the precipitate Schnee ‘snow'	199
59	Distribution and aggregates frequencies of the NON-SPATIAL modifier variable across the three constructions (ARG+fallen, ARG+geben, and es+PRED) with the precipitate Schnee ‘snow'	201
60	Raw and relative frequency of the GENRE variable in the ARG+fallen construction with the precipitate Schneeregen ‘sleet’ (abbreviated as SR)	207
61	Raw and relative frequency of the TAM variable in the ARG+fallen construction with the preciptate Schneeregen ‘sleet’	208
62	Raw and relative frequency of the SPATIAL modifier variable in the ARG+fallen construction and the preciptate Schneeregen ‘sleet’	209
63	Raw and relative frequencies of the NON-SPATIAL modifier variable in the ARG+fallen construction with the preciptate Schneeregen ‘sleet’	211
64	Frequencies of different sentence positions of es ‘it’ in concordance lines returned for queries (xxi) and (xxii)	217
65	Raw and relative frequencies of the GENRE variable within the ARG+geben construction and preciptate Schneeregen ‘sleet’ (SR)	217
66	Raw and relative frequencies of the TAM variable in the ARG+geben construction and the precipitate Schneeregen ‘sleet’ (SR)	218
67	Raw and relative frequency of the SPATIAL modifier variable in the ARG+geben construction and the preciptate Schneeregen ‘sleet’	220
68	Raw and relative frequencies of the NON-SPATIAL modifier variable across the ARG+geben construction with the preciptate Schneeregen ‘sleet’ (SR)	221
69	Frequency of conjunctions for nominal co-MEs of the precipitate Schneeregen ‘sleet’ by TAM factor	224
70	Raw and relative requencies for the two variables TAM and GENRE for the es+PRED-Cx with the precipitate Schneeregen ‘sleet’	226
71	Raw and relative frequency of the SPATIAL modifier variable in the es+PRED-Cx and the preciptate Schneeregen ‘sleet’	228
72	Raw and relative frequencies for the NON-SPATIAL modifier variable within the es+PRED-Cx and the precipitate Schneeregen ‘sleet’	229
73	Raw and relative frequencies of the GENRE variable across the three constructions (ARG+fallen, ARG+geben, es+PRED) and precipitate Schneeregen ‘sleet’, (abbreviated as SR)	231
74	Raw and relative frequencies of the TAM variable across the three constructions (ARG-fallen, ARG-geben, es+PRED) and the precipitate Schneeregen ‘sleet’ (abbreviated as SR)	233
75	Raw and relative frequencies of the SPATIAL modifier variable across the three constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Schneeregen ‘sleet’ (abbreviated as SR)	235
76	Distribution and absolute frequencies of variable MODIFIER across the three constructions (ARG+fallen, ARG+geben, es+PRED) and precipitate Schneeregen ‘sleet’ (abbreviated as SR)	237
77	Frequencies of literal and figurative USAGE of the ARG+fallen construction with the precipitate Hagel ‘hail’	242

78	Raw and relative frequencies of the GENRE variable for the ARG+fallen-Cx and the precipitate Hagel 'hail'	243
79	Raw and relative frequencies of the TAM variable for the ARG+fallen-Cx and the precipitate Hagel 'hail'	245
80	Raw and relative frequencies of the SPATIAL modifier variable within the ARG+fallen-Cx and the precipitate Hagel 'hail'	246
81	Raw and relative frequencies of the GENRE variable for the ARG+fallen-Cx and the precipitate Hagel 'hail'	248
82	Frequencies of literal and figurative USE for the ARG+geben-Cx and the precipitate Hagel 'hail'	253
83	Raw and relative frequencies of the GENRE variable for the ARG+geben-Cx and the precipitate Hagel 'hail'	254
84	Raw and relative frequencies of the TAM variable for the ARG+geben-Cx and the precipitate Hagel 'hail'	255
85	Raw and relative frequencies of the SPATIAL modifier variable for the ARG+geben-Cx and the precipitate Hagel 'hail'	257
86	Raw and relative frequencies of NON-SPATIAL modifiers for the ARG+geben-Cx and the precipitate Hagel 'hail'	258
87	Distribution and frequencies of the precipitate Hagel 'hail' as a co-ME in combination with the preposition mit 'with' and the conjunction als 'as' within the returns of the ARG-Cxx	262
88	Frequencies of literal and figurative USAGE for the es+PRED-Cx and the precipitate Hagel 'hail'	263
89	Raw and relative frequencies of the GENRE variable in the es+PRED-Cx and the precipitate Hagel 'hail'	263
90	Raw frequencies of the TAM variable within literal instances of the es+PRED-Cx and the precipitate Hagel 'hail'	264
91	Frequencies of the SPATIAL modifier variable within the literal es+PRED-Cx and the precipitate Hagel 'hail'	266
92	Distribution and frequencies of the NON-SPATIAL modifier variable for the es+PRED-Cx and the precipitate Hagel 'hail'	267
93	Frequencies of literal and figurative USE by construction type (ARG+fallen, ARG+geben, and es+PRED) with the precipitate Hagel 'hail'	269
94	Raw and relative frequencies of the GENRE variable by construction type (ARG+fallen, ARG+geben, and es+PRED) and the precipitate Hagel 'hail'	270
95	Raw and relative frequencies of the tam variable across all three constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Hagel 'hail' (literal instances only)	273
96	Raw and relative frequencies of the SPATIAL modifier variable across all three constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Hagel 'hail'	274
97	Raw and relative frequencies of types of NON-SPATIAL modifiers across the three constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Hagel 'hail'	276
98	Frequencies of cognate constructions by precipitate in the deTenTen13 corpus and the DWDS corpora	299
99	LogDice scores and most frequent subjects collocating with verb rieseln in the deTenTen13 and in the DWDS Wortprofil	318

100	Frequencies of the Regen/Schnee rieseln collocations in the deTenTen13 and the DWDS corpora for PREC2 and PREC1 word order	322
101	Frequencies of deTenTen13 search queries containing Frau Holle and the Schnee 'snow' precipitate	332
102	Frequencies of the am- and beim-progressives in the deTenTen13 and the DWDS Metakorpora	341
103	Frequencies of preposition bei ‘at’ plus noun precipitates in the deTenTen13 and DWDS Metakorpora	346
104	Definitions and examples of the preposition bei ‘at/by’ adopted from the DWDS dictionary	347
105	Frequencies of bei+precipitate (conditional and simultaneous events) out of 100 concordances per precipitate in the deTenTen13 corpus	350
106	Frequencies of German "bad weather" compounds in the deTenTen13 and the DWDS corpora and their acceptance by Duden online	360
107	Frequencies of different compounds of precipitate plus -wetter 'weather' and the corresponding prepositional construction bei …-wetter in the deTenTen13 corpus and the DWDS Metakorpora	362
108	Frequencies of different compounds with -wetter 'weather' and the corresponding prepositional construction bei …-wetter in the deTenTen13 corpus and the DWDS Metakorpora	364
109	Frequencies of figurative use per precipitate and construction type based on the coded data frames downloaded from the deTenTen13 corpus examined in Chapter 3	367

List of Figures

1	The scalar typology of p-encoding (adapted from Eriksen et al. 2010: 593)	16
2	The symbolic structure of a construction (Croft 2001: 18)	37
3	Example output of CQL search for syntax [word="es"][word="Regen"][lemma="geben"] sorted to the left context	70
4	DWDS word cloud of typical collocations of the noun Schnee 'snow'	87
5	Visual representation of frequencies of the lemmata Schneeregen 'sleet' (left) and schneeregnen 'to sleet' (right) in some of the corpora of the DWDS	89
6	Screenshot of the first four returns for the search query Schneeregen && fallen in the DWDS WebXL corpus	90
7	Distribution and absolute frequencies of the three constructions (ARG+fallen, ARG+geben, es+PRED) within the Regen 'rain' precipitate data across the GENRE variable	148
8	Relative percentual frequencies of construction type (ARG+fallen, ARG+geben, es+PRED) with the precipitate Regen 'rain' per type of GENRE	149
9	Absolute frequencies and distribution of TAM factors across the three constructions (ARG+fallen, ARG+geben, es+PRED) with the precipitate Regen 'rain' and with categories pres_FUT, FUT, and IRR subsumed under FUT_intent ("future intent")	151
10	Relative frequencies in percentage of each individual construction (ARG+fallen, ARG+geben, es+PRED) per total of TAM category within the precipitate Regen 'rain' data sets, present_FUTURE, FUTURE, and IRREALIS are subsumed under FUTURE_intent	152
11	Distribution and absolute frequencies of SPATIAL MODIFIERS across the three constructions (ARG+fallen, ARG+geben, es+PRED) with the precipitate Regen 'rain'	155
12	Absolute frequencies of non-spatial modifiers across the three constructions (ARG+fallen, ARG+geben, es+PRED) with the precipitate Regen 'rain'	158
13	Percentual proportion of each individual construction (ARG+fallen, ARG+geben, es+PRED) per total of NON-SPATIAL modifier within the precipitate Regen 'rain' data sets	158
14	Absolute frequencies of the GENRE variable across the three constructions (ARG+fallen, ARG+geben, es+PRED) in combination with the precipitate Schnee 'snow' ("other" and figurative instances removed)	193
15	Relative frequencies in percentages of the GENRE variable across constructions (ARG+fallen, ARG+geben, es+PRED) with the precipitate Schnee 'snow'	194
16	Absolute frequencies of construction (ARG+fallen, ARG+geben, es+PRED) by TAM variable in combination with the precipitate Schnee 'snow' (levels pres_FUT, FUT, and IRR subsumed under "FUTURE_intent"; FIGURATIVE/OTHER removed)	196
17	Relative contribution of each construction (ARG+fallen, ARG+geben, es+PRED) to level of tam for the Schnee 'snow' precipitate data in percentages	197
18	Overall distribution and frequencies of SPATIAL modifiers across the three constructions (ARG+fallen, ARG+geben, and es+PRED) with the precipitate Schnee 'snow'	199
19	Relative contribution of each construction (ARG+fallen, ARG+geben, es+PRED) to each spatial modifier type for the Schnee 'snow' precipitate data in percentages	200

20	Distribution and absolute frequencies of the NON-SPATIAL modifier variable across the three constructions (ARG+fallen, ARG+geben, and es+PRED) with the precipitate Schnee ‘snow'	201
21	Proportional contribution in percentages of construction per level of NON-SPATIAL modifier within the Schnee ‘snow’ precipitate data sets	202
22	Distribution and absolute frequencies of the GENRE variable across the three constructions (ARG-fallen, ARG-geben, es+PRED) and the precipitate Schneeregen ‘sleet’	232
23	Distribution of absolute frequencies of the TAM variable across the three constructions (ARG-fallen, ARG-geben, es+PRED) and the precipitate Schneeregen ‘sleet’	233
24	Relative frequencies of the TAM variable for the precipitate Schneeregen ‘sleet’ by construction (ARG+fallen, ARG+geben, es+PRED) with levels present_ FUTURE, FUTURE, and IRREALIS collapsed under FUTURE_intent	234
25	Distribution and absolute frequencies of the SPATIAL modification variable across the three constructions (ARG-fallen, ARG-geben, es+PRED) and precipitate Schneeregen ‘sleet’	235
26	Distribution and absolute frequencies of the NON-SPATIAL modifier variable across the three constructions (ARG-fallen, ARG-geben, es+PRED) with the precipitate Schneeregen ‘sleet’	237
27	Percentual proportion each construction (ARG-fallen, ARG-geben, es+PRED) contributes to the total of each level of the NON-SPATIAL modifier variable with the precipitate Schneeregen ‘sleet’	238
28	Distribution and absolute frequencies of the GENRE variable across all three constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Hagel ‘hail’	271
29	Percentual proportion of all three constructions (ARG+fallen, ARG+geben, es+PRED) and the Hagel ‘hail’ precipitate to the individual type of GENRE (with “other” removed due to the low incidence)	271
30	Distribution and absolute frequencies of levels of the TAM variable across all three constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Hagel ‘hail’	273
31	Percentage of each individual construction (ARG+fallen, ARG+geben, es+PRED) per total of TAM level within the precipitate Hagel ‘hail' returns	274
32	Distribution and relative frequencies of the SPATIAL modifier variable across the three constructions (ARG+fallen, ARG+geben, es+PRED) and the Hagel ‘hail’ precipitate	272
33	Absolute frequencies of the NON-SPATIAL modifier types TEMPORAL, QUANTITY, QUALITY, and STANCE across all three constructions constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Hagel ‘hail’	277
34	Relative frequencies of the NON-SPATIAL modifier types TEMPORAL, QUANTITY, QUALITY, and STANCE across all three constructions constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Hagel ‘hail’	278
35	Relative frequency of construction type by types of GENRE in percentages	287
36	Relative frequency of construction type by TAM level in percentages	287
37	Relative frequency of construction type by SPATIAL modifier in percentages	288
38	Relative frequency of construction type by NON-SPATIAL modifier in percentages	288
39	Relative frequency of precipitate type by type of GENRE in percentages	289
40	Relative frequency of precipitate type by TAM level in percentages	290

41	Relative frequency of precipitation type by SPATIAL modifier in percentages	291
42	Relative frequency of precipitate type by NON-SPATIAL modifier in percentages	293
43	Frequency curve of the multi-word sequence Leise rieselt der Schnee 'Softly falls the snow' in the Google N-gram viewer between the years 1800–2000 with an inset of the years 1900–1920	317
44	Word trajectory of the term Fritz-Walter-Wetter in the DWDS newspaper corpus from 1946–2016 (downloaded from the DWDS)	357
45	Exercise sheet lexical field "The weather" for beginning learners of German	389
46	Exercise sheet lexical field "The weather" for intermediate learners of German	391

CHAPTER 1

Stormy Weather Expressions and Constructions

"*The rain began again. It fell heavily, easily, with no meaning or intention but the fulfilment of its own nature, which was to fall and fall*." (Garner 1992: 150, *Cosmo Cosmolino*)

1.1. Introduction

The quote above, an observation of a weather event, seems fairly commonplace. There appears to be nothing remarkable in terms of the depiction of rain, but doubtlessly, the novelist would have chosen her words with utmost deliberation. Despite the fact that the rain has no bearing whatsoever on the storyline, the quote is special enough to be featured on a number of quotation websites and it can even be bought online, on a laser-engraved a wooden plaque. We are thus presented with the interesting juxtaposition of a mundane event in the guise of an evidently extraordinary, quote-worthy surface expression or, rather, several expressions relating to the same event. Consequently, the duality of Garner's quote represents a microcosm of how we speak about the weather. While the topic is ostensibly simplistic, linguistically, the expression of weather events is anything but. In fact, the properties of meteorological predicates are peculiar (Ruwet 1991), to say the least, and not just because their subjects do not behave like prototypical subjects do. Syntactically, weather expressions vary remarkably. This is true not just across different languages, but also within a single language, which can display considerable diversity in how weather events are expressed.

Matters of literary style aside and following typological accounts to be discussed below, the most natural or neutral way to speak about precipitation events in a language like English or German is by means of a so-called "**PREDICATE TYPE**" construction, as established by Eriksen, Kittilä and Kolehmainen (2010). Their investigation has partly prompted the study at hand. According to them, in Germanic languages the predicate is chiefly responsible for encoding the meteorological event while the subject does not tend to refer to the weather event at all (or anything else, for that matter, being a so-called "dummy" constituent or "expletive"). To stay within our typological boundaries, we should therefore expect a sentence such as *It started raining* or *It started to rain again*, instead of "The rain began again" as the first part of the quote

above. In the quoted case, the subject is responsible for encoding the fact of the meteorological event while the predicate does not reference the weather event directly. The second part of the quote touches on other issues that lie at the very heart of the present study. The author—unwittingly—presents us with a wealth of linguistic topics which will be examined throughout this dissertation.

Let us briefly consider some of the issues raised by the quote that are extensively discussed in later chapters. As expressed in the first sentence, rather than using a predicate-based expression, the author chose a nominal argument, *rain*, in combination with a (light) predicate, *fall*. This construction is not just an example of what Eriksen et al. (2010) call the "**ARGUMENT TYPE**" of weather construction, but the verb *fall* itself affords us a glimpse into the linguistic compression of motion events (more thoroughly discussed in Section 1.2). The specification of a PATH ('down') is conflated with the fact of non-intentional motion. Subsequently, there is the reduplication of *to fall and fall* which introduces us to important modifying material, marking intensification or continuation or, here, simply describing what rain *does*—which is furthermore present in the adverbs *heavily* and *easily*, providing us with quality and/or manner. The rather philosophical phrase "with no meaning or intention but the fulfilment of its own nature" is testament to the non-agentive nature of precipitation—which brings us back to the non-prototypicality of the subjects of weather events.

Prototypical subjects, by contrast, are not just agentive but animate and definite (cf. Keenan 1976: 319; Comrie 1989: 189; Langacker 1991: 305–309; Dixon 1994: 8, 119). These attributes are particularly evident in transitive clauses where a prototypical subject contrasts with a prototypical object/patient, which is defined by being inanimate and indefinite (see Hopper and Thompson 1980; Rice 1987; and Comrie 1989: 189). It is no surprise, then, that in most languages, weather events come in intransitive packaging. The focus of mainstream research in syntax has commonly been on argument structure constructions in transitive and ditransitive clauses. By contrast, very little attention has been paid to the simplest clause type of all, the intransitive clause. In traditional and theoretical approaches to grammar, the intransitive clause is generally perceived of as very straightforward and basic in nature, since, from a syntactic viewpoint, an intransitive verb only takes a single argument. Moravcsik, for example, gives the following definition of intransitive: "a verb that needs no object; for example, *exist*" (2006: 242). The discussion usually ends there. Intransitives seem to be considered as linguistic one-acts or one-man shows, requiring the most minimal set of marking demands in a language.

However, even a one-act isn't confined to a single scene and the intransitive clause is not without complexities. On the basis of weather expressions alone, the present study demonstrates that there is more to the lowly intransitive clause than traditional or formal grammars will have us believe; that this purportedly simple one-act is at the least a complex one, if not, in fact, a full-length play. I aim to provide a construction-based analysis of meteorological events in German. As such, this study is a departure from the majority of previous studies that focus wholly on syntactic properties of weather constructions. This dissertation will rely on an approach that is not exclusively syntactic, but also lexico-semantic in nature. Moreover, it is based on fully contextualized data representing actual usage, as extracted from the German web corpus, the *deTenTen13*, available through the corpus manager and text analysis software site, *Sketch Engine* (hereafter *SE*).

Subscribing to the philosophical commitments of cognitive linguistics (notably, Lakoff 1987, 1991; Langacker 1987, 1990, 1991; and Talmy 2000), I do not assume clear-cut boundaries between different levels of linguistic analysis such as lexicon, morphosyntax, and semantics/pragmatics, but rather consider these to be on a representational if not processing continuum. Following this assumption, semantics cannot be divorced from syntax. Intransitives should not be disregarded as unworthy of in-depth analysis based on the presupposition that a syntactically minimal argument set is uninteresting semantically.

The choice of weather events as a vehicle for exploring some of the research questions laid out above is a natural one considering that the weather is, without a doubt, a tried-and-true conversationally safe topic that anyone can be affected by and have an opinion on as all of us are affected by the weather. More importantly, it can be the gateway to other topics, providing a means of building rapport and establishing a connection with an interlocutor. It thus has an important social function. Considering the universality of weather events, their pervasiveness, and their importance for everyday human life, both in the physical world and as phatic lubricant, it is dismaying that there are only a few linguistic analyses of expressions of weather events in the literature (cf. Darden 1973; Ruwet 1991; Paykin 2010; Bleotu 2012, 2013; Meulleman and Stockman 2013; Eriksen, Kittilä and Kolehmainen 2010, 2012, 2015; Kolehmainen 2008, 2010; Andrason 2019, Andrason and Visser 2017, 2019; Huang, Dong, Yang & Ren 2021, Liégeois 2021). Moreover, the available studies largely tend to focus on argument structure (cf. Meulleman and Paykin 2016) or are typological in nature, thus skimming along a wide surface of phenomena instead of going in-depth. By contrast, this dissertation goes deep and goes into

detail by zooming in on one language, German, and does so with extensive corpus data. As mentioned above, the cross-linguistic variability of weather expressions, as demonstrated by their morphosyntactic vicissitude, is intriguing. While examples from all sorts of weather events are adduced to illustrate the complexities of weather expressions, the data presented here largely focus on meteorological events involving precipitation, such as rain, snow, sleet, and hail.

The structure of this introductory chapter is as follows. Section 1.2 gives additional context about research on intransitives, specifically Leonard Talmy's seminal work on motion events (which present a classic intransitive clause type if ever there was one) and the components that go into the marking of such events. Section 1.3 reviews the semantic expression and construal of weather events and provides a synopsis of past research on weather constructions. The discussion of previous research, furthermore, introduces several issues involving weather expressions. One of these concerns the actual "substance" of weather arguments, i.e., the "stuff" (Ruwet 1991: §3) that precipitation itself and other weather types are made of. Section 1.4 reveals the special status of the non-prototypical and even problematic subjects of weather events, such as the expletive *es* 'it'. Section 1.5 addresses questions of frameworks and terminology taken in this study. Specifically, I give an overview of the fundamentals of construction grammars and cognitive linguistic approaches as well as their shared assumptions, while untangling the interconnectedness between these two overarching frameworks. Moreover, it elaborates on usage-based linguistics, which subsumes both construction grammars and cognitive linguistics (although usage-based approaches, in principle, are not tied to any specific grammatical framework). I focus especially on the importance of usage-based data and factors to consider when choosing a corpus-based approach. Section 1.6 comes full circle and discusses the fairly recent field of lexical typology of which Talmy's study on motion events was in the vanguard and now is considered a classic typological study of the lexicon. The structure of the remainder of this dissertation is presented in Section 1.7.

1.2 Intransitives and Talmy's typology of motion events

As a rule, in traditional accounts of syntactic structure, the focus has been on multi-participant constructions or impersonal constructions, but not on intransitives per se.

Intransitives are usually dismissed in discussion not exceeding a single paragraph, as exemplified by the following passage from the *Duden*, the canonical German reference dictionary, first published in 1880 and considered to be the authority on the "correct" use of the German language:

> Intransitive Verben (Intransitiva)
>
> Intransitiv werden Verben genannt, die eine Subjektleerstelle eröffnen und dieser eine semantische Rolle zuordnen, aber kein Akkusativobjekt – auch kein reflexives – fordern. Je nach der Verbbedeutung kann das Subjekt eine Agens- oder eine Patiensrolle tragen: *lachen*, *weinen*, *wandern*, *denken*, *trauern*; *sterben*, *fallen*, *liegen*, *stehen*[1]
>
> (*Duden* 2006: Section 540, 403–404)

By contrast, in cognitive approaches, there tends to be more attention paid to the form and meaning of the elements in intransitive clauses, in general, and in weather expressions, in particular. Langacker, for example, discusses "non-distinct argument phenomena" (1991: Section 8.3.), which often surface as fully intransitive clauses. As opposed to prototypical transitive clauses where the participants are clearly discernible from one another and other entities, non-distinct argument phenomena include those where process and participant are not clearly delimited and differentiated; this is especially true for meteorological events (hereafter MEs). The notable exception with regard to the paucity of studies of the intransitive is Leonard Talmy's research, which has examined and openly theorized about classic, prototypical intransitives in the guise of motion events (1972, 1975, 1985, 1991, 2000).

Talmy developed a cognitive semantic framework that looks at how a particular type of intransitive clause, a non-caused motion event, is encoded. This research has spawned a large program of cognitive linguistics research investigating how languages structure space (cf. Levinson 2003 for an overview).[2] A motion event is a situation "containing movement or maintenance of a stationary location" by a single participant (Talmy 1985: 61). Talmy identified a number of semantic components pertaining to motion events, based on the assumption that these components can be isolated conceptually, but not necessarily linguistically. In a motion event, there are four central components, FIGURE, GROUND, PATH and MOTION, and two optional

[1] 'Intransitive verbs are verbs that open up an empty subject position which they assign a semantic role which needs neither an accusative object, nor a reflexive one. Depending on the meaning of the verb, the subject can be an agent or a patient: *laugh*, *cry*, *wander*, *think*, *grieve*; *die*, *fall*, *lie*, *stand*' (my own translation).

[2] Talmy, of course, also examines caused motion events.

co-events, MANNER and CAUSE (Talmy 2000: 25). These semantic components (in this dissertation, I use small caps for recurring semantic concepts) are expressed by a variety of syntactic elements or "surface components" (Talmy 1972: 57), but there is largely not a one-to-one relationship between them. A semantic element can be expressed by more than one surface element, but a combination of semantic elements can also be expressed by a single surface element. The following examples demonstrate how the verb root (in bold) conflates LOCATION + MANNER in (1), self-agentive MOTION + MANNER in (2), non-agentive MOTION + CAUSE in (3), and non-agentive MOTION + MANNER in (4), while the PATH is expressed in the prepositional phrases (underscored). In these examples, the FIGURE is (1) *the vase*, (2) *Rocky*, (3) *a plastic bag*, and (4) *the ball*, respectively. The GROUND is (1) *the table*, (2) *the steps at the Philadelphia Museum of Art*, (3) *the beach*, and (4) *the stairs*, respectively.

(1) *The vase **stood** on the table.*
(2) *Rocky **ran** up the steps at the Philadelphia Museum of Art.*
(3) *A plastic bag **blew** across the beach.*
(4) *The ball **bounced** down the stairs.*

These examples may not seem very intriguing to a speaker of English, but we find that in other languages, speakers will be hard-pressed to encode the very same sentences in the same way that English does, as described in the following. Talmy's seminal work gave rise to a major two-way typology of languages based on how languages encode the basic semantic component of PATH in a motion event. Correspondingly, Talmy introduced the notions of verb-framed (or V-languages) and satellite-framed languages (or S-languages).[3] V-languages express the PATH component of a motion event in the verb root itself. S-languages, to which English belongs, express PATH in a so-called satellite. The term satellite covers several grammatical elements across languages. In English, PATH is mainly expressed in prepositions and adverbial particles (which greatly overlap with prepositions), as we have seen in (1)–(4). In other satellite-framed languages like Russian and German, the PATH satellite can take the shape of prefixes (***hinein****rennen* 'run **into',** (***he***)***runter****fallen* **'**fall **down'**).

[3] Talmy, in fact, suggested more than just these two types of motion predications, but verb-framed and satellite-framed languages seem to be the dominant types (at least with regard to the more well-described languages). Thus, minor types, such as MOTION +FIGURE conflating languages (e.g., the Hokan language Atsugewi; or Athapaskan, cf. Rice 2002) are often simply not considered in the discussion.

Talmy's typology is not without its detractors and some of the critiques are addressed below. To give just one example which showcases how problematic "true" translations can in fact be, consider an assumedly simple English sentence such as *She danced into the hall*. As an S-language, the fact of MOTION and the MANNER of movement are conflated together and expressed in the verb root (*dance*), while the PATH is expressed via the satellite (*into*). It is not possible, though, to impose this seemingly simple structure with the very same meaning onto a verb-framed language such as Spanish, as the PATH component tends to be contained in the verb root. MANNER of MOTION, if salient enough to warrant its expression, must be expressed in a separate phrase or constituent. Thus, in Spanish, we can only say *Ella bailó **a** la sala* 'She danced (up) **to** the room' or *Ella bailó **en** la sala* 'She danced **in**(side) the room'. The meaning of the latter can only be locational and not directional. If we want to express the PATH with the end location (*into*), we are stuck with the MANNER-poor *Ella **entró** a la sala* 'She **entered** the room'. The MANNER component can then be expressed separately, often in the shape of a gerund, as in *Ella entró a la sala*, ***bailando*** 'She entered the room, **dancing**'. In narratives, Spanish speakers will often make up for the lack of MANNER richness in their motion lexicon by putting more focus on the context, e.g., by describing the setting to a greater extent (cf. Slobin 1996).

Talmy did not explicitly discuss intransitives, but his investigations have shed light on how a variety of semantic components have a bearing on event realization in intransitive clauses. These components include motion, agency, causation, valence, and information structure. Intransitives may not offer a wealth of semantic content, but because they have been so neglected descriptively and theoretically, it is intriguing to investigate everything that goes into and interacts with an intransitive (e.g., lexical choice, tense-aspect-mode [TAM] marking on the verb, negation, semantic prosody, stance, preferred subject, and modification).

With regard to the dichotomy in Talmy's typology, it has become apparent upon closer inspection that it does not hold up. Counterexamples are aplenty and problematic for Talmy's typology. For example, putative verb-framed languages, like Spanish and Japanese, do indeed allow for MANNER verbs in PATH frames in some cases, as shown in (5) and (6).

(5)	*La botella*	***flotó***	*hacia*	*la*	*cueva*.	[Spanish; Aske 1989]
	the bottle	floated	up to	the	cave	
	'The bottle floated to the cave.'					
(6)	*John-wa*	*kishi-made*	***oyoida/tadayotta***.			[Japanese; Beavers 2008]
	John-TOP	shore-until	swam/drifted			
	'John swam/drifted to the shore.'					

Aske (1989) has argued that a satellite-framed pattern in a verb-framed language is permissible when the PATH is atelic, but not with telic paths that are resultative. Slobin, who has contributed greatly to the typology of motion events in various publications, has refined Aske's suggestion and introduced the "boundary-crossing constraint" (Slobin and Hoiting 1994; Slobin 1997):

> It appears to be a universal characteristic of V-languages that crossing a boundary is conceived of as a change of state, and that state changes require an independent predicate in such languages. ….When a path crosses a boundary, then, it is no longer possible to accumulate a series of grounds to a single verb, because the state-change from one side of the boundary to the other will be expressed by a separate verb with its associated ground (e.g., Slobin 1997: 441).

Moreover, Slobin (2004: 228) suggests assigning languages with bipartite verbs (Algonquian, Hokan, and Athapaskan language families [cf. Rice 2002, for the latter]), serial verbs (cf. Zlatev and Yangklang 2004 for Thai), and MANNER and PATH preverbs (cf. Schultze-Berndt 2007 for Jaminjung) to a third class, that of "equipollently-framed languages", as both PATH and MANNER have to be expressed in one or more verbs that conflate MANNER and MOVER. The verb complex, then, comprises both MANNER and PATH information in two (or more) obligatory main verbs or verb stems. These studies have led to the conclusion that languages, instead of being confined to a two-way typology, may be better conceived of as placed on a continuum (Slobin 2004).[4] However, the mere fact that languages cannot simply be labelled as one of two types with respect to motion-event framing, in this case either verb- or satellite-framed, does not take away from Talmy's other findings. Talmy's research has been extremely influential and continues to spawn a plethora of research on motion events with an end yet to come.

Lastly, not only is Talmy's the only research that has taken intransitives seriously, but it also constitutes one of the seminal studies regarding typological research on the lexicon. The emerging field of lexical typology is discussed in more detail in Section 1.6 The next section reviews the linguistic literature on weather events and the different perspectives that analyses

[4] Talmy himself points out, for example, that English (and German, for that matter) has instances of FIGURE+MOTION conflation as in the non-agentive *It rained in through the bedroom* window, where the FIGURE of rain is conflated with the fact of MOTION, or the agentive *He spat into the cuspidor*, where the implicit FIGURE of spit is moving (2000: 96). This is the same for the German equivalent *spucken* 'spit', but also for *würfeln* 'roll a die', which in English does not conflate the figure of the die with the motion event.

have taken.

1.3 Prior analyses of weather constructions

Studies on the weather in the linguistic literature are still slim, but recently, the topic has gained increased and deserving attention. More research has been published, which approaches these complex events from different angles. This section provides an overview of some of the previous analyses of weather constructions germane to the dissertation at hand, especially those that typologize the weather in a manner similar to Talmy's approach to motion events. Studies that have been mentioned in the opening pages of this manuscript but are not further discussed here are Kolehmainen (2008, 2010), Paykin (2010), Meulleman and Stockman (2013), Eriksen et al. (2015), Levin and Krejci (2019), and Huang et al. (2021). These studies generally focus on aspects of weather event expressions, which have either little or no bearing on the study at hand, or do not contribute new information, such as kinesis, prototypicality of the subject, or unaccusativity. This is also true for Ruwet's (1991) second chapter on weather, namely *Weather Verbs and the Unaccusativity Hypothesis*. Therefore, I focus instead on Ruwet's first weather chapter *On Weather Expressions*. Darden (1973) and Bolinger (1977) are mentioned in passing as they are discussed by Ruwet.

Regarding the structure of this literature review, this section is bookended by the research of two of Talmy's contemporaries: Ray Jackendoff and Nicolas Ruwet. Jackendoff's (1983) considerations only afford one paragraph to the conceptual structure of "rain". Ruwet (1991), on the other hand, bestows two whole chapters of weather analysis on the reader in *Syntax and Human Experience*, which more aptly should be named "Human Experience and Syntax". He discusses many issues regarding mostly French weather expressions, one of which involves the "non-referentiality" of the expletive subject including a discussion of the "substance" of precipitation, which I elaborate on in more detail while adding my own deliberations on German at the end of this section. The first three studies—Jackendoff (1983), Bleotu (2012, 2013), and Meulleman and Paykin (2016)—relate to Talmy's work in that they make the connection between weather events and motion events, but otherwise, they have little bearing on the study at hand. Meulleman and Paykin (2016) take up Jackendoff's analysis. Their contribution is discussed briefly. This is followed by Bleotu (2012, 2013) who analyzes weather verbs as motion verbs and focuses on a possible argument structure representation for

these verbs. Her research is mentioned because some of her suggested subjects include examples pertinent to the German language.

These accounts are followed by the studies which are the most germane to this dissertation. Other than Ruwet, the most in-depth discussion of previous literature is afforded to Eriksen et al.'s cross-linguistic examination of motion events (2010 and 2012), which partly prompted the dissertation at hand. They typologize a number of languages and thus, stay on the surface. By way of contrast, this surface-level analysis triggered my own deep and comprehensive examination of German precipitation events. In a similar fashion, Andrason (2019) and Andrason and Visser (2019) draw on Eriksen et al.'s typology for a more in-depth discussion of Polish and the South African language, isiXhosa. They raise similar issues with the typology as I do for German. Before I return to Ruwet's (1991) contribution, I discuss Liégeois's (2021) recent study, which does not relate to the above mentioned papers, but revolves around the text genre of German weather reports, which are a genre that I examine as well (cf. Chapter 3). Thus, this work ties in nicely with the variables I identified and my findings. He provides empirical data on how weather reports are communicated to a specialised audience (such as meteorologists) versus a lay audience (the public) by focusing on the weather phenomena themselves with special regard to the time and place of their occurrence. Last but not least, Ruwet's (1991) discussion of weather events is concerned with the syntax-semantics interface, and he makes many compelling points. Accordingly, his work is fundamental and referenced by many, if not all, of the above-mentioned studies. In the following, the authors of the studies in question are bolded at the beginning of each review for easier readability.

Jackendoff (1983) and **Meulleman and Paykin (2016).** In a manner similar to Talmy, Jackendoff (1983) suggests that semantic structure is conceptual structure, but only insofar as it serves the syntax of a language. He deconstructs several lexical verbs, such as *to butter* and *to dust,* but also *to rain,* and analyzes the latter as having the conceptual structure given in (7):

(7) [Event GO ([Thing RAIN], [Path DOWNWARD])] (Jackendoff 1983: 185)

Here, both the Thing ('rain') and the Path ('downward') are constituents of the underlying or schematically construed Event ('go'). *To rain* is the only meteorological verb that Jackendoff analyzes by semantic decomposition (in fact, it is the only precipitation verb he includes). Talmy (whose 1985 work Jackendoff references among others), however, analyzes the verb *rain* as an

instance of the minor FIGURE+ MOTION pattern (see fn.4), where the FIGURE of rain is conflated with the fact of motion, which is rare in the English language (2000: 57, cf. fn.4). Jackendoff and Talmy's analyses share common ground in that FIGURE and Thing are viewed as the same thing. Where they differ is that there is no Path explicated in Talmy's analysis. Moreover, Jackendoff's principal concern is the conceptual structure of specific English verbs for eventual syntactic realization, while Talmy's is typologizing languages based on the lexicalization patterns they exhibit in motion events (cf. Meulleman and Paykin 2016). Path is a conceptual component, so, even though Talmy ultimately classified languages, his classification is based on differential patterns of conceptual structure as observed in typologically distinct languages. As Meulleman and Paykin themselves point out on morpho-lexicological grounds, "it is not quite clear whether its [=*to rain*] conceptual structure should be taken as limited to these two constituent elements [FIGURE+ MOTION]" (2016: 61). I, for one, do not think so. Obviously, due to our knowledge of the world, it is understood that the trajectory of rain (or precipitation in general) must ultimately be downwards, even if it is "raining sideways". Therefore, the Path is implied and does not have to be specified. This is comparable to Spanish speakers being able to say *Un pájaro entra a la casa* 'A bird enters the house'—where MOTION + PATH are conflated—without having to specify the MANNER component, as flying is the default type of locomotion associated with birds. Thus, while the Path is not overtly lexicalized in *to rain*, it does not mean that it is not part of a speaker's conceptual structure. Compare an inverse case from Finnish, where the precipitation verb *sataa* 'to rain' originally meant simply 'to fall', but it has grammaticalized to refer exclusively to precipitation exclusively in its literal usage, although the Thing or FIGURE is not overtly encoded, only implied (Eriksen et al. 2010: 588–589).

Meulleman and Paykin (2016) discuss the semantic and syntactic status of weather verbs, based on Talmy's (2000) and Jackendoff's (1983) analyses. Their main focus lies in the consideration of weather verbs as motion verbs and how Talmy's dichotomy (2000), as discussed in Section 1.2, relates to this consideration as well as to the homo- vs. hetero-geneity of weather verbs as a class. They come to the conclusion that weather verbs do not form a homogeneous class. It should be noted that Talmy himself gives *rain* as a FIGURE+ MOTION conflating example in the English language, but he does not claim that precipitation verbs form one homogenous class. Analogously to Jackendoff's example (repeated below in [7']) Meulleman and Paykin give their own example (see (8)), showcasing the conceptual structure of *thunder/lightning* to drive home their point about weather verbs not reducing them to a special

class of motion verbs.

(7') [Event GO ([Thing RAIN], [Path DOWNWARD])]

(8) [Event APPEAR ([Thing NOISE/LIGHT], [Path DISCHARGE TRAJECTORY])]

In their conclusion, they reiterate the heterogeneity of weather verbs in terms of their conceptual structure and in relation to the PATH component, their dynamicity, and also the different substances involved in expressions of weather. Lastly, they point out the "systematic incorporation of MANNER" (2016: 66) as a core component of weather verbs and how Talmy's (2000) binary typology does not account for their distribution across languages. This assessment is based on their analysis of weather verbs as motion verbs.

Bleotu (2012 and 2013). Bleotu (2012, 2013), too, analyzes weather verbs as motion verbs. Nevertheless, taking a more syntactic perspective, she is mainly concerned with unaccusativity or unergativity of weather verbs, which has no bearing on the study at hand. In following Hale and Keyser's (2002) conflation theory of verb formation, she presents a possible argument structure representation based on cross-linguistic data (Germanic and Romance languages, Latin, and Chinese). In doing so, she not only looks at precipitation verbs, but also gives a few examples dealing with atmospheric MEs (e.g., Spanish *está oscuro* 'it is dark' and *hay niebla* 'it is foggy'). Bleotu claims, from a small sample of languages, that "weather verbs (…) lack a real subject" and that they may take any of the following subjects (examples are taken from Bleotu 2012: 61 and 2013: 48f.)

(i) pro-drop (e.g., Spanish/Italian): *Llueve.* 'rains'

(ii) an expletive (e.g., German/English): ***It*** *is raining.*

(iii) a demonstrative (e.g., colloquial German/Dutch)
Das *regnet schon wieder!*
this rains already again
'It is already raining again.'
(This is possible in Dutch with *het* or *dat*, both 'it', where the latter is emphatic)

(iv) a personal pronoun (e.g., Icelandic)
Hann *er farinn adh rigna.*
he is started to rain
'Oh, sh*t, it's raining again!'

(v) a noun denoting the background or the source of the weather phenomenon (e.g., *nature*, *sky*, *God*)
a. *that cloud… / the sky… / God thunders*
b. or the substance of the weather phenomenon itself:
The rain is raining on us heavily.

Example (v) b. basically corresponds to Eriksen et al.'s (2010) **ARGUMENT-PREDICATE COGNATE TYPE**, where the expression of the ME is split between the argument and the verb as in a cognate construction, such as *rain is raining*. One basic claim that Bleotu also makes is that weather verbs are motion verbs. Following Hale and Keyser (2002), she suggests that *rain* can be decomposed as V+N (rain= 'FALL RAIN') for unaccusatives and as V+N ('CAUSE ['FALL RAIN]') for unergatives. (2013: 47). The reasoning underlying this dichotomy is based on typological data. First of all, languages like Chinese have no weather verbs. Instead, Chinese uses a combination of *fall* (v.) + *rain* (n.) to express precipitation. Secondly, the Finnish and Hungarian verbs meaning *to precipitate* originally meant *to fall*. This generalized meaning has been lost and the verbs (Finnish *sataa* and Hungarian impersonal *esik*) now exclusively refer to precipitation (Bleotu 2012: fn.7). Andrason points out that the same development is true for Polish *padać* 'fall, rain', whose original sense is 'fall' (2019: 100). Although not necessarily transparent anymore, 'to fall' would still be part of the conceptual structure. Lastly, languages are able to paraphrase, alternating between unaccusatives and unergatives (e.g., *es regnete* 'it rained' vs. *Regen fiel* 'rain fell'). In addition to the last claim about paraphrasing, Bleotu acknowledges that languages may follow different strategies in encoding precipitation events, even if they actually possess weather verbs. She lists the following constructions (apart from V+N like French *tombe la pluie* 'rain falls'):

(vi) impersonal constructions

a. *Piove* (Italian) lit. 'Rains'; 'It rains.'

(vii) extraposed 'subject' constructions

b. *Il a plu toute la journée une petite pluie fine.* (French)
EXPL has rained all.FEM.SG the day a.FEM.SG little.FEM.SG rain smooth.FEM.SG
'It rained a heavy rain.'

(viii) agent constructions

c. *The Lord thundered from heaven.*

d. *He rained his tears on me.*

In her 2013 paper, Bleotu also discusses the semantic lightness or heaviness of verbs that appear with weather events (as an example, *be* and *come* are considered lighter than *give* and *fall*) and proposes that the "'heavier' light verbs (be decomposed) into primitive predicates" (Bleotu 2013: 47)

Other lexical frameworks which have concerned themselves with the question of unergativity or unaccusativity of weather verbs include: Ruwet (1991); Benincà and Cinque (1992); Paykin (2010); and Meulleman and Stockman (2013). These considerations have no bearing on the dissertation at hand.

Eriksen, Kittilä, and Kolehmainen (2010 and 2012). The study that partly triggered this dissertation is Eriksen, Kittilä, and Kolehmainen's (2010 and 2012) cross-linguistic examination of meteorological events (hereafter MEs), which bears a family resemblance to Talmy's approach to motion constructions. Accordingly, their contribution to the study of weather events is given the most space in this overview. Eriksen et al.'s (2010, 2012) research concerns the question of how languages deal with encoding events that have no "proper" or prototypically construed participants. Their aim is, first and foremost, to suggest a formal typology of ME constructions in general, and eventually a more specialized typology of languages, with a focus on precipitation events (cf. 2010: 567). Their exploratory effort leads to a major three-way classification of ME constructions, including several sub-types, based on whether the argument, the predicate, or a combination of both is responsible for coding the specific weather type in a weather expression. Their Section 3 (Eriksen et al. 2010: 571–586), which makes up the bulk of their study, focuses on the dominant construction types that are employed cross-linguistically for expressing weather. This typology includes all types of weather participants, such as *sun*, *wind*, and *temperature*—and not just precipitation events. For the

sake of completeness, their illustration of the three major types of construction types is reproduced in Table 1 with a supplementary summary of each type. The most important takeaway is that German is classified as a predicate-type language of the expletive subtype.

Table 1. *Schematic illustration of the typology of meteorological expressions (adapted from Eriksen et al. 2010: Figure 1; [summaries mine])*

Language type	Subtypes	Description
PREDICATE TYPE	**VALENCY VARIATION** Atransitive type Expletive type Intransitive predicate type Transitive predicate type **PARTS OF SPEECH VARIATION** Verbal type Adjectival type Adverbial type Nominal type Locative type	**SUMMARY** The predicate (verbal or non-verbal) mainly encodes the ME. Predicate types are subcategorized according to syntactic valency of the predicate and the part of speech of the predicate.
ARGUMENT-PREDICATE TYPE	Cognate type Split type	Predicate and argument either equally express the same information (as in cognates: *rain rains*) or predicate and argument express different attributes of the same event (as in split types: *the wind blew*).
ARGUMENT- TYPE	Intransitive argument type Existential type Transitive argument type	The ME is primarily encoded in the argument (*rain*, *snow*, *hail,* etc.) which is syntactically realized as the subject of a transitive or intransitive construction. The predicate in argument type constructions is deemed semantically light or vacuous (*fall*, *go*, and so on).

Based on their typology of possible ME construction types, Eriksen et al. proceed to suggest a tentative formal typology of languages (2010: 586–596), as presented in Figure 1. The prototypical scalar typology is based on how precipitation is encoded (and precipitation only, or what they call *p-encoding*), as opposed to other environmental expression events, such as temperature, which are nevertheless part of the all-encompassing typology of ME constructions in Table 1. They maintain that the main distinction is between argument p-encoding and the options to the right in Figure 1, with argument p-encoding and predicate p-encoding representing "extreme oppositions" (2010: 593). The typology can also be read as a function of

the predicate's semantic involvement in the precipitation encoding, with the richest or strongest predicate semantics to the left and the weakest to the right of the scale. Conversely, the argument's semantic involvement is strongest to the right of the scale and weakest or non-existent to the left. Languages do not necessarily engage in one type of p-encoding only, it should be noted.

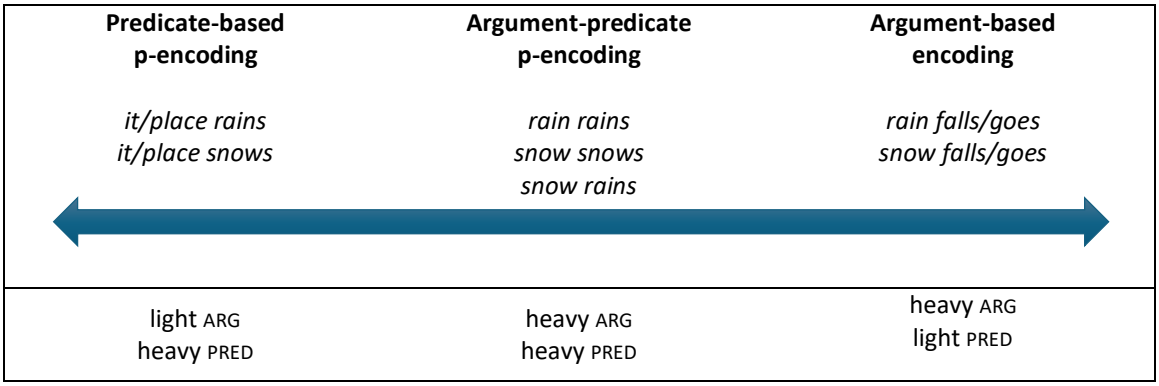

Figure 1. *The scalar typology of p-encoding (adapted from Eriksen et al. 2010: 593)*

Eriksen et al. (2010) assert that languages may manifest themselves as intermediate types. Such an example is Turkish, which allegedly lies between argument and generalized p-encoding, such that both argument and predicate contribute semantically within the typology. However, they claim that Germanic, Romance, and Sami languages consistently display predicate p-encoding. This typology places German on the far left of the scale as a predicate p-encoding language, or "expletive type", a type that according to Eriksen et al. "is very rare outside Europe" (2010: 574). At the same time, they maintain that this cross-linguistically infrequent ME encoding type "illustrate(s) classic examples of the use of expletives" (2010: 573). Eriksen et al. state that "the key for language variation in MEs is whether languages allow the predicate type" since "the existence of the argument type is a trivial fact of languages" (2010: 587). This claim hinges on the premise that all languages allow for NPs that denote more "tangible participants" like *rain*, *hail*, *sun*, and so on, or at least participant-like entities of MEs (as discussed in Section 3, event and participant—or entity—cannot always be easily teased apart). Therefore, they believe that every language must at least possess the argument type to encode precipitation events. Even languages such as German that predominantly employ the predicate type still have ample options to utilize the argument type, as illustrated in (9), where the argument is bolded and underscored.

(9) a. ***Regentropfen*** *peitschten gegen mein Fenster.*
‘**Raindrops** were lashing against my window.’
b. *Kalter* ***Regen*** *strömte die Schornsteine hinunter.*
‘Cold **rain** was pouring down the chimneys.’

Eriksen et al.’s observation about “the triviality of the argument type” holds only for precipitation events, but not for other environmental comment events such as temperature, ambience, and noise (2010: 588). This statement is based on their reasoning that precipitation involves at least quasi-participants (such as *snow*[*flakes*] and *rain*[*drops*]), while temperature and noise are virtually intangible, and make for even less of a (quasi-)participant. Eriksen et al. acknowledge that a language may employ more than one construction type for coding weather events. Russian, for example, uses the argument type only with regard to p-encoding, with a heavy argument and a light verb such as *idti* ‘go’, as in (10). Conversely, temperature is encoded using the predicate type, as in (11). They conclude that precipitation events naturally lend themselves to argument-type, but not predicate-type coding. Thus, any language encoding a precipitation event will at least have the argument type, while languages predominantly using a type other than the argument type, will have the argument type at their disposal as well.

(10) *Idet* *sneg.* [Russian]
go.3SG.PRES snow
‘It is snowing.’

(11) *Kholodno.* [Russian]
cold.SG.N.PRED
‘It is cold’.
(cf. Eriksen et al. 2010: 586, their example [49b])

Eriksen et al.’s typology is very much grounded in syntax. However, they also discuss the semantic contributions of the predicate and argument in a construction, as I have annotated at the bottom of Figure 1. They do not believe that the supportive predicate in the argument type language is particularly empty semantically (or what I am calling “light”). Consider the verb *fall*, for example, which conflates MOTION and PATH; while not empty semantically, it is far lighter than a verb like *gush* ‘to flow rapidly and copiously’. Conversely, they claim that a predicate-type language is able to leave the argument position truly empty semantically in the guise of an expletive subject or no subject at all (cf. 2010: 587–588). Apart from this assessment, though, they do not go into detail about what this assumed semantic contribution looks like or what

roles the respective items play in a weather construction. If we consider the semantic contribution of each item necessary to encode the ME, in line with Eriksen et al.'s observations, we find that not everything is equally available. For example, not every ME specifies the setting. Consider the German cognate type in (12), taken from an Austrian novel, where both argument and predicate, redundantly, make the same semantic contribution. The precipitate appears in both elements; it is neither an unspecified *it* that is raining, nor a semantically heavy argument that is complemented by a light verb such as 'fall'.

(12) ***Der Regen regnete*** *auf jenen Stein, auf jenen nicht.* (Thomas Glavinic: *Die Arbeit der Nacht*)[5]
'**The rain rained** on that stone, but not on that one.'

In a nutshell, the semantic contribution or heaviness of argument and predicate obviously differs. In the argument p-encoding constructions, the argument does the semantic heavy lifting. In the predicate p-encoding constructions, the predicate takes on that role. The cognate type, *rain rains*, most evenly shares the semantic load between argument and predicate, where both are semantically "heavy" and neither of the elements is backgrounded. The split type, *snow rains*, is similar with an equal division of "heavy-heavy", but predicate and argument are expressing different attributes of the same event. It is doubtful, though, that there is a language that would employ the split type to encode something like *rain snows* as a standard option. In neither the cognate nor split type expression is the setting acknowledged (unless it is expressed in a separate constituent), while in the predicate type, such as in German, it is blended into and represented by the pronoun *es* 'it'.

In terms of the notions of heavy and light, the argument type does not always have to adhere to the "heavy argument + light predicate" assumption. Looking at clauses like the one in (13), both argument and predicate are heavy (the latter could even be considered super-heavy), with the precipitate encoded in the argument and the predicate conflating MOTION and MANNER.[6]

[5] Source: *DWDS* contemporary corpus (novel: 'Night Work', 2006)

[6] In a way, *stürzen* also conflates PATH, as it means 'to fall in a bad/violent way'. *Falling* generally implies downward motion. It is possible to use *stürzen* with an *upward* or *outward* direction; however, in such cases, the direction has to be expressed additionally in a satellite as in *Er stürzte zur Tür* ***hinein***/***hinaus*** 'He bolted **in**/**out** the door' or *Er stürzte die Treppe* ***hinauf*** 'He rushed **up** the stairs'. The direction is furthermore reinforced by the preposition *vom/vom Himmel* 'from/from the sky'. Moreover, we can easily add the adverbial *herab* 'down' or *hinunter* 'down': *Regen stürzt vom Himmel herab/hinunter* 'Rain is gushing down from the sky'.

(13) ***Regen stürzt*** *vom Himmel.*
'**Rain is gushing** from the sky.'

On the other hand, there are "light + light + heavy" combinations such as *es gibt Regen* lit.'it gives rain' or *es fällt Schnee* 'it falls snow' where we find three constituents: (1) a setting/participant amalgamation in the expletive pronoun, (2) a light predicate, and (3) a heavy precipitate noun. In case of the latter with snow, there are even two possibilities: *es fällt Schnee* 'it falls snow' and *Schnee fällt* 'snow is falling' where, depending on the word order, the *es* 'it' is possible or ungrammatical.[7] The *es gibt* 'it gives' construction, on the other hand, is fixed and it is not possible to dispose of the *es* 'it'. Categorizing these "light + light + heavy" constructions as intransitive MEs might be generous, but they are not full-blown transitives either.

What exactly Eriksen et al.'s typology is based on is not fully transparent. According to their own assertion, they extracted their data largely from their own knowledge, reference grammars, and a "sample of MEs in 25 languages" obtained from questionnaires (2010: 567). The content of the questionnaires is not available. Moreover, there are no frequencies given, only the disclaimer that their sample size is too small to have statistical relevance. Unsurprisingly, their examination skims a number of available constructions and languages but does not penetrate the surface. Optional adverbs and adverbials, for example, are not part of Eriksen et al.'s considerations (2010: 571). In light of the fact that their study is an admittedly exploratory effort, perhaps any degree of analytical depth cannot be expected. In similar fashion to Talmy's assessment of German as "verb-framed", they call German "the predicate type", and possibly erroneously so. Although German displays the "expletive type" of ME-encoding as frequent or as the most basic option (according to reference grammars), as illustrated in this dissertation, there is a richness of available and frequently used ME constructions in German that has been ignored in Eriksen et al.'s examination. This holds true not only for the types of constructions that are possible (i.e., not just the predicate type), but also for modifications of the basic construction types by optional elements that speakers often exploit. Disregarding the variety of constructions that MEs get packaged in means disregarding speakers and their

[7] *Im Winter ist es häufig frostig und* ***es fällt Schnee***. 'In the winter it is often chilly and **it falls snow.'** (Source: context.reverso.net, last accessed May 30, 2024) or ***Es fällt Schnee*** *in Südafrika*. '**It falls snow** in South Africa.' (Source: de.nachrichten.yahoo.com, last accessed May 30, 2024). Ungrammatical: ***'…, *wenn es Schnee fällt* 'when it snow falls' (but: …. *wenn es schneit* ***or*** *wenn Schnee fällt* '… when it is snowing' or '… when snow is falling' are grammatical').

knowledge and use of a language. It may be superficially descriptive, but it is neither justified nor useful to reduce German to the "expletive type", when there is an abundance of ME constructions available for specific purposes. Part of the purpose of this dissertation is to explore actual usage conventions in German ME encoding and to advocate that the emergent field of lexical typology, discussed in Section 1.6, not make the same mistake as traditional typology in making strong overarching claims from a paucity of data. In the remainder of this dissertation, I use the abbreviations *es*+PRED-Cxx, ARG-PRED-Cxx and ARG-Cxx respectively, to refer to the different construction types corresponding to the three ME types that Eriksen et al. (2010) proposed.

Andrason (2019) and **Andrason and Visser (2019).** More recently, and based on Eriksen et al.'s (2010, 2012) typology, Andrason (2019) offers a qualitative, comprehensive description of Polish based on introspection and acceptability ratings from native speakers. He states that he interviewed native speakers, but does not specify how many, how he chose the participants, or in which manner these interviews were conducted. He points out, however, that elicited utterances, which he uses as examples in his paper, were counterchecked against various written texts (2019: 71). Andrason covers not just precipitation events but makes an effort to address "dynamic events", which can involve precipitation or non-precipitation, such as thunder, lightning, and wind, as well as "static events", which are either atmospheric, such as temperature, humidity, and light, or involve light emission, such as sunshine. This dichotomy of dynamic and static events is based on Eriksen et al. (2012: §3.1). Andrason finds that ME constructions in Polish are very complex and offer "a slightly distinct profile" each, depending on the ME (2019: 102). He states that his examination of Polish corroborates some of Eriksen et al.'s hypotheses, but also notes that his more in-depth analysis of Polish reveals greater complexity than provided by standard grammars and Eriksen et al.'s typology, which mirrors my own findings for German. For example, he claims that "each (meteorological) phenomenon can be expressed by a wide variety of constructions" and that "the encodings of any two (MEs) diverge" (2019: 100, 102). Accordingly, he found that the dynamic encodings "are usually compatible with all types of encoding" (2019: 98), but mostly fall back on the intransitive argument-predicate type or the atransitive predicate type. He suggests that other possible subcategories may be relevant, e.g., constructions with peripheral instrumental objects (2019: 102). Andrason, moreover, notes usage differences between precipitate types—a finding that has crystallized for the German data I present in subsequent chapters. For example, he claims

that the concept of 'rain' allows for the widest range of constructions and lexical items that encode the ME. Moreover, there is no translation equivalent for 'sleet', for which a paraphrase involving snow and rain (or vice versa) must be employed (but not a compound as in German *Schneeregen*, lit. 'snowrain'). Furthermore, he asserts the paucity of the 'hail' precipitate occurring in the predicate constructions. This finding, too, appears to echo the German data, which have revealed that the *es*+PRED-Cx (*es hagelt* 'it is hailing') is predominantly used for figurative utterances (cf. Chapters 3 and 4 below). Additionally, Andrason encourages the exploration of items like 'good weather' (*pogoda*, which also means just 'weather'). This item may be similar to German compounds explored in Section 4.3.3 (e.g., *Kaiserwetter* 'emperor weather' for 'beautiful weather'); however, Andrason does not list "other" types of 'weather', such as a concept for 'bad weather', just *pogoda* 'good weather' (2019: 92–93).

Andrason and Visser (2019) come to similar conclusions following Erikson et al'.s (2010, 2012) framework in a description of isiXhosa, a Bantu language indigenous to southern Africa and one of the official languages of South Africa and Zimbabwe. Here, too, 'rain' events appear in all three major construction types (the predicate type [PRED-Cxx], the argument type [ARG-Cxx], and the argument-predicate type [ARG-PRED-Cxx]) while the other three precipitate types ('snow', 'sleet', and 'hail') are more restricted in their expression, lexically and syntactically. According to Andrason and Visser, "there is no verb that could specifically and exclusively denote snow, hail, or sleet" (2019: 24) where, furthermore, the words for the latter two could also refer to specific realizations of *rain*, e.g., a thick or cold rain, or even snow. Neither the 'hail' nor the 'sleet' precipitate occurs in predicate constructions either, but in ARG-Cxx and ARG-PRED-Cxx constructions only, while the 'snow' event is even more restricted and only occurs in ARG-Cxx. Again, these findings lightly reflect the German data examined throughout this dissertation in that the *Regen* 'rain' precipitate appears to be the most flexible and prolific with regard to the syntactic and lexical variation it allows, while the other three precipitate types appear to be more restricted or, at least, less frequent in general. Furthermore, Andrason and Visser (2019) observe an onomatopoeic source for verbal and nominal precipitation types, which is also true for some German rain verbs (e.g., *pladdern* 'patter' or *plästern*, roughly 'to pour', as briefly explored below in Section 4.4. Lastly, Andrason and Visser (ibid.) also concerned themselves with the grammaticalization path of weather verbs with cognate objects in six African languages from South Africa.

Liégeois (2021). Another recent study, by Liégeois (2021), resembles the study at hand in how it analyses genres of weather events, i.e., weather reports, and the modifiers under scrutiny. More specifically, Liégeois examines how weather reports are communicated in texts addressing lay audiences (e.g., the public) versus expert interlocutors (e.g., meteorologists). His study is restricted to the three domains constitutive of weather reports, which are 'weather situation' ("*Sachobjekt* [= *Wetterlage*]") or what I refer to as *weather* or *meteorological event*), *Raum* 'space', and *Zeit* 'time' as identified by Brinker, Cölfen, and Pappert (2018: 48). These three elements have furthermore been established as being central for WEATHER frames, by both *Berkeley Frame Net* and *Frame Net des Deutschen* 'German Frame Net' at the *University of Düsseldorf*.[8] Obviously, a weather forecast must be specific concerning these three elements, but a more informal "weather talk" situation may not be as consistent in spelling out all three: precipitate type, space/place, and time. In addition to the precipitation type, the *Berkeley Frame Net* frame for PRECIPITATION also determines PLACE and TIME as core elements, but it also specifies non-core elements, such as DURATION and RATE or QUANTITY, EXPLANATION, FREQUENCY, MANNER, and TEMPERATURE.[9]

Liégeois found that a nominal style dominates both of the text types that he analyzed (2021: 226f.), i.e., phrases indicating the *days of the week* or *time of the day* are more frequent than adjectives or adverbs where temporality is under consideration. The *time of the day* is the most frequent identifier in both text types (with more than 48% and 41%, respectively), while items specifying the *day of the week*, e.g., *am Donnerstag* 'on Thursday', were much more frequent in the texts addressing lay audiences (approximately 30.5% vs. almost 19%). The third most frequent temporal modifier overall is *deixis,* which is fairly evenly distributed between the two text types (19.3% in expert texts vs. 17.9% in lay texts). Liégeois does not define his classifications, but merely gives examples like "Deiktika (z.B. *dann*, *immer*)" '"deictics" (then, always)'. Less frequent categories such as *Dauer* 'duration' and *Abfolge* 'sequence' make up a vanishingly slight percentage of his data with reference to lay audiences (0.85% combined) and 4.2% combined with reference to expert audiences. Lastly, Liégeois found that "weather" and

[8] Source: *Berkeley Frame Net* at https://framenet2.icsi.berkeley.edu/fnReports/data/luIndex.xml (last accessed May 24, 2024); *German Frame Net* https://framenet-constructicon.hhu.de/framenet (last accessed May 24, 2024)

[9] *Berkeley Frame Net* at framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Precipitation (last accessed January 10, 2023).

"location" are more frequent items than "time". He concludes that temporal items are not as frequent due to the fact that weather reports follow a chronologically linearized structure where the same point or period in time is linked to one or more weather phenomena in one or more forecast regions (2021: 225). We shall keep in mind his classification of temporal items, which are germane to this dissertation, but that his data consist exclusively of "weather reports", albeit different text types. Liégeois's examination of locatives or *Raum* 'space' has little bearing here, as the terminology or jargon used in texts addressing expert audiences appears to have little overlap with the findings of this dissertation. For example, there is marginal overlap with his categories of *Windrichtung* 'wind direction' and *Deixis*, but items referring to the layers of the Earth's atmosphere, like "trophosphere", are few and far between, so that they do not warrant their own category in the data at hand.

Table 2 is adapted from Liégeois and shows his classification of temporal items as well as the frequency of both lemmata and tokens in texts prepared for expert audiences (e.g., meteorologists, left column) vs. texts prepared for lay audiences (e.g., the general public), right column).

Table 2. *Classification and frequency of temporal items in texts prepared for an expert audience vs. a lay audience (adapted from Liégeois 2021: 226.f.)*

Rank	Expert audience	Lemmata	Tokens	Lay audience	Lemmata	Tokens
1.	**time of day**	41	471	**time of day**	18	338
2.	**deixis**	14	188	**day of week**	7	251
3.	**day of week**	7	185	**deixis**	7	147
4.	**sequence**	11	89	**sequence**	6	76
5.	**duration**	6	24	**other**	3	6
6.	**other**	13	17	**duration**	1	1

These items identified by Liégeois play a role in Chapter 3, where I comprehensively analyse the three major construction types ([*es gibt*+ ARG_{PREC}], [*fallen*+ARG_{PREC}], [*es*+V_{PREC}][10] with the respective precipitate type (*Regen* 'rain', *Schnee* 'snow', *Schneeregen* 'sleet', and *Hagel* 'hail') and the modifiers they tend to occur with, such as LOCATION, TIME, QUANTITY, QUALITY, and STANCE. The temporal modifiers are obviously the ones most pertinent to Liégeois's study.

Ruwet (1991). Last but not least, one of the most compelling studies in the linguistic literature

[10] Subscript $_{PREC}$ refers to precipitation in particular whereas subscript $_{ME}$ includes other types of MEs not involving precipitation.

on MEs is without question Ruwet's (1991) account in *Syntax and Human Experience*, which takes a slightly different approach to examining weather expressions. Ruwet discusses "problems lying on the border between syntax and semantics" (1991: xvii). Despite the fact that he was not a cognitive linguist, his study of MEs can be characterized as taking a nongenerative and meaning-based approach. He concerns himself with the relation between meaning and grammar, trying to make sense of a semantically homogeneous verb class (weather verbs) that appears in not-so-homogeneous syntactic trappings. In acknowledging the structural variety of weather expressions cross-linguistically, he suggests that they may not lend themselves to being described by the traditional grammatical categories of subject and predicate, recognizing "a conflict between our experience and the analytic demands of syntax" (1991: 103). Ruwet largely examines French weather expressions (but also brings in examples from Russian, Japanese, Basque, Georgian, and Arabic), and considers, in particular, whether the "expletive", when present, is genuinely an expletive. Ruwet admits that his findings are not conclusive and need more in-depth study. He assumes, for example, a difference between the uses of French *il* (as in *il pleut* 'it rains') and *ça* (as in *ça pleut* 'that rains') which are in competition with one another and appear in different stylistic contexts. *Ça* is colloquial and lacks an English counterpart in weather expressions (**that is raining*); nevertheless, its prevalence in French merits analysis in its own right, he argues.

The French *ça*-construction seems to mirror the German *das regnet* 'that is raining' vs. *es regnet* 'it is raining' (also see Bleotu's (iii) demonstrative example above). In some varieties of German, it is possible, and customary, to say *das regnet* 'that is raining'. In northern Germany, in particular, *das regnet* is used interchangeably with *es regnet*. In other geographical regions, this option is less common but also valid, especially when modified emphatically, as in (14) and (15). This type of stanced and intensified construction, for example, is not addressed by Eriksen et al. (2010, 2012) at all.

(14) ***Das*** *regnet (ja) schon wieder!*
'**That** is raining yet again!'

(15) ***Das*** *regnet (ja) immer noch!*
'**That** is still raining!'

In both (14) and (15), the particle *ja* expresses astonishment or surprise or, possibly, indignation. These are just two examples that show how a speaker can modify the expression of an ME,

communicating not just stylistic and regional differences, but also attitudes.

Ruwet makes numerous intriguing observations and raises excellent points in his chapter, *On Weather Expressions* (1991: Ch.3), many of which are echoed in this study on German MEs. Many of his insights revolve around the problem of the referentiality of the "null or expletive" subject in weather expressions (particularly in French and how the French syntax differs from the English syntax with regard to weather expressions). This is important as he eventually focuses on Bolinger's (1977) and Darden's (1973) accounts, which are largely based on English. Ruwet's conclusion is that there is a conflict between the analytic demands of syntax and our immediate experience or our "pretheoretic representation (*Vorstellung*)" (1991: 103). In line with this, the peculiarity of weather expressions simply mirrors the peculiarity of the corresponding events and their mostly diffuse agents whose aspects are partially expressed in the predicate. For that reason, sentences like those in (16) and (17) are unacceptable, as in our experience they are "subsisting beings (…) whose existence is for us largely independent of the growth of the grass (itself barely perceptible), (or) of the fall of the stone …" (1991: 104)

(16) **Il herbe.*
'It grasses', or, 'it's grassing'

(17) **Il a pierré.*
'It stoned.'

At least in English and German, though, while *grasses* or *grassing* may not be possible, *grünen* 'to green', certainly is. The German meaning seems to be slightly more expansive than in English, as it includes the meanings 'to sprout/to spring/to spear'. Thus, it is not only grass that can 'green'. The key lyric, *the rain in Spain stays mainly in the plain*, from the musical *My Fair Lady*, in the German version translates to *es grünt so grün, wenn Spaniens Blüten blühen* 'it greens so green, when Spain's blossoms blossom. *Es grünt so grün* 'it greens so green' is a commonly encountered expression in German (together with the collocation *es grünt und blüht* 'it greens and blooms').

While English and French apparently lack impersonal verbs when it comes to diurnal expressions, in German it can indeed *tagen* 'to day', *nachten* 'to night', and *dämmern* 'to dawn' or 'to dusk'.[11] The former two are used more regionally (e.g., Switzerland) though, or in more

[11] German makes no difference between *dusk* and *dawn*, at least in the predicate. As a noun, of course there is the possibility to distinguish the compounds *Morgendämmerung* 'morning …' and

elevated style. Interestingly, English has kept *dawn* as a verb, but it does not *dusk* anymore, as it did in Middle English (*dusken* 'to become dark'). And, apparently, Old English had a whole paradigm of these verbs: *it days, it evens, it winters*, etc. (cf. Ruwet 1991: 119). Neither has Modern English *dawn* become a "two-in-one", meaning the same surface form with different but related meanings, as in the German *dämmern.* Ruwet gives the two examples, *the day is dawning* and (*the*) *dawn is breaking*, to highlight how, even within one language, "the division of the experience into subject and predicate may shift". (1991: 105). Examples like these tie in nicely with the findings of this dissertation. Eriksen et al. (2010, 2012) claim that languages predominantly use one construction type. What needs to be stressed here is the word "predominantly". As soon as we start digging a bit deeper, we find that there is a multitude of ME constructions for a speaker to choose from, based not just on the precipitation type, but on a range of factors, explored in the following chapters. It may well be true that a language resorts to a default construction, but it is obvious that ME constructions are not evenly distributed cross-linguistically or even within a language and, as I will show, they are not evenly distributed lexically or inflectionally either. Ruwet's work recognizes and highlights this important finding, but Eriksen et al.'s (2010, 2012) proposed typology neither captures the depth of a language's repertoire nor does it address the differences showcased by Ruwet and described for German throughout this dissertation.

One interesting point that Ruwet makes is his discussion of the "substance" of precipitation (or what he also refers to as "stuff"), more specifically between *rain* and *snow* (1991: §3). This consideration is of importance to the present study with regard to many of the utterances involving *Schnee* 'snow', since it is often difficult to tease apart what a speaker is referring to: the actual precipitation event of snow that is in the process of falling or snow that is on the ground? Ruwet claims that snow is a "mass" and, thus, a "stuff". As opposed to *rain*, *thunder*, or *lightning*, which are treated as individuals in French (cf. *rain falls* vs. ***la*** *pluie tombe*), Ruwet considers *snow* to be ambiguous. *Snow* can be the same "dispersed" FIGURE as *rain*, but it can also be a "stuff", much like *sand* or *mud* (1991: 101). He quotes Mantha and Mel'čuk (1984:

Abenddämmerung 'evening …'. The *Duden Online* dictionary gives one of the meanings of *Dämmerung* as *Halbdunkel*, lit. 'half dark'. Thus, a German speaker can use *dämmern* figuratively in a similar manner to the English *It dawned on me*. *Na, dämmert's?* is a phrase used to colloquially express something like 'Do you get it now?'. It is interesting that the English version makes more sense in this metaphorical use, as it is 'the dark' that is lifted or gradually lifting, whereas in German it could be both (*dusk*/*dawn*), but obviously only the meaning of 'to dawn' is mapped. *Dämmern* (v.), figuratively, cannot mean that something that is clear becomes more obfuscated or unclear.

312–313) who say that "rain is … a substance (=water) … but rain is a substance that cannot be seen at rest, unlike snow". Ruwet interprets this to mean that *rain* only denotes a precipitation but not a "stuff". He exemplifies how *a walk in the snow is ambiguous* whereas *a walk in the rain* can only be one thing. I whole-heartedly agree with Mantha and Mel'čuk's second part of the quote. Doubtlessly, the fact that *snow*, as opposed to *rain*, can and does linger influences a speaker's use of ME constructions and the frequency of different tenses used in the constructions (cf. Chapter 3 of this dissertation). I disagree with Ruwet's interpretation, though, that "*rain* does not denote a substance: the substance of rain is water". The latter part of his statement is true for the substance of *snow* as well since *rain* and *snow* are both water in a different state of aggregation. Ruwet tries to corroborate this point by reminding us that we can "eat snow" but cannot "drink rain", which I take to mean that we cannot drink rain unless it is collected in a container. Interestingly, German has the compound *Schneewasser* 'snow water', meaning snow that has thawed and turned into water, which also has its own entry in the *Duden*.

An additional attempt at exposing the difference of substance of the precipitates is given by Ruwet by examining how natural the nouns *rain* and *snow* sound when used without further modification in ARG-Cxx. He states that the French noun *neige* 'snow' is not completely natural in denoting precipitation and lists the French examples, as well as two English equivalents for consideration; these are given in (18) and (19) (taken from Ruwet 1991: 102). The French 'snow' examples seem at least questionable.

(18) a. *pendant la pluie* *'?during the rain(fall)'*
b. *la pluie a cessé* *'the rain stopped'*
c. *the rain lasted two hours*

(19) a. *??pendant la neige* *'?during the snow'*
b. *??la neige a cessé* *'the snow stopped'*
c. *??the snow lasted two hours*

Ruwet accepts the following two examples given by Mantha and Mel'čuk (1984: 292) but points out that he would prefer *rainwater* instead of *rain* in (d) below.

d. *On a installé de grands réservoirs pour recueillir la pluie.*
'We have installed large tanks to gather rain.'

e. *Il secoua* [sic] *la pluie qui alourdissait son chapeau.*[12]
'He shook the rain which weighted down his hat.'

With regard to German, while the noun *Schnee* 'snow' seems a little less natural than *Regen* 'rain' in similar sentences, there is no shortage of examples for both. A search on the German Google domain for the exact words *während des Schnees* 'during the.$_{GEN}$ snow.$_{GEN}$', which implies snowfall, returned 88,400 results.[13] Many of the images accompanying the texts, do indeed show people or animals in falling snow. The returns for *Regen abschütteln* 'shake off rain' (N=201) have a higher incidence than for *Regenwasser abschütteln* 'shake off rainwater' (N=60).[14] By contrast, a search for *Schnee abschütteln* 'shake off snow' returned 6,730 results. Some of these combinations are exemplified in (20) and (21). Example (21) also contains the precipitate type *Graupel* 'graupel' (also 'snow pellets' or 'soft hail', cf. merriam-webster.com), where the latter, in following Ruwet, would constitute more of a "stuff" at rest than *rain*.

(20) *Barry Clemons (…) betrat den Laden und* ***schüttelte Regen und Graupel von seinem Mantel****.*[15]
'Barry Clemons (...) entered the store and shook **rain and graupel** from his coat.'

(21) *Thut eure schweren Bündel und Ledersacke ab,* ***schüttelt euch den Schnee von Kleidern*** *und Füßen, so gut es gehen mag, und folgt mir ins Zimmer.* (Canisius, 1867: 128)
'Take off your heavy bundles and leathersacks, **shake the snow from the clothes** and feet as well as possible and follow me into the room.'

Clearly, in German, it is possible to shake off *graupel*, *rain* and *snow*, although the latter appears to occur more frequently. In (22) the precipitation noun must be reasonably interpreted as *falling snow* rather than *snow on the ground* due to the construction with the verb *stop* (the same would be true for *start*).

---

[12] This example appears to be based on the sentence *Il neigeait, et Rabe secoua la neige qui alourdissait son chapeau* 'It snowed, and Rabe shook off the snow that weighed down his hat' from Pierre Mac Orlan's novel *La Quai des Brumes*, published in 1926 (Source: https://storage.gra.cloud.ovh.net/v1/AUTH_3877b9485d06455c8b7c11a9fadaeb2d/attachments/original/4/8/4/002622484.pdf last accessed May 30, 2024).

[13] Search carried out on May 30, 2024.

[14] Also, a quick search (the exact words in quotation marks) for *Regen auffangen* 'collect rain' on the German *Google* domain returns fewer results (N=16,000) than for *Regenwasser auffangen* 'collect rainwater' (N=49,700).

[15] www.ylva-verlag.de/leseprobe/leseprobe_beziehung-ausgeschlossen-von-jae.pdf (last accessed November 12, 2019)

(22) *Auch am Abend* ***hörte der Schnee noch nicht auf***. (german.china.org.cn)
'Even in the evening **the snow did not stop**.'

In contrast, a sentence like that given in (23), without the modifier *fallend* 'falling' in parenthesis, would indeed be ambiguous (cf. Ruwet's example *a walk in the snow*) as to whether the speaker is referring to 'falling snow' or 'snow-covered terrain'. (The use of the preposition *durch* 'through' with MEs is briefly discussed in Section 4.3.2).

(23) *Wir wanderten zwei Stunden* ***durch den (fallenden) Schnee***.
'We hiked **through the (falling) snow** for two hours.'

Ruwet points out that French has the verb *enneiger* 'to cover with snow' just like there is *ensabler* 'to cover with sand', but that there is no parallel construction for **empleuvoir* 'to cover something with rain'. Similarly, German has the following prefixed verbs involving *Schnee* 'snow':

i. ***verschneien*** 'completely cover with snow';
ii ***einschneien*** 'cover with snow' or 'snowed in, so that you cannot leave a place'; and
iii. ***zuschneien*** 'to be covered with/filled in with/blocked by snow'[16]

and the following prefixed verbs involving *Regen* 'rain' and *Hagel* 'hail':

iv. ***verregnen***, meaning that something has been 'spoiled by rain'.[17]
v. ***verhageln***, which translates to 'spoil' or 'ruin' and is almost always used figuratively.[18]

Additionally, the *Duden* gives two different variants for *einregnen*, a reflexive impersonal version and a plain personal version with two subsenses:

vi. ***sich einregnen*** (reflexive, impersonal), meaning 'to not stop raining / to become an incessant rain'
vii. a. ***einregnen*** meaning 'to be drenched with rain'
b. ***einregnen*** meaning 'rained in so that you cannot leave a place' (which parallels

[16] Definitions as per the *Duden Online* dictionary (https://www.duden.de/woerterbuch).

[17] For example, *Unser Mittagsspaziergang ist absolut* ***verregnet****, wobei der Begriff Regen noch geschönt ist.* 'Our lunchtime walk is absolutely **spoiled by rain**, where the term rain is certainly euphemized.' (Source: *deTenTen13*, lenaundmerlin.cipa.de)

[18] For example*, Sehr sehr schade, uns* ***verhagelt*** *dieses Votum richtig die Wochenendstimmung!* 'It's a real shame, this vote really **ruins** our weekend mood!' (Source: *deTenTen13*, ms-reporter.de)

einschneien (in ii).[19]

This section has discussed some of the previous literature on weather events in general as well as the matter of the "substance" of tangible precipitates. While I have already alluded to the problematic nature of the arguments involved, the next section focuses on the role of the so-called "expletive" subject, particularly how it relates to weather events and its "non-participant status" or ostensible "meaninglessness".

1.4 Below the intransitive: What is weathering?

In line with Hopper and Thompson's (1980) and Rice's (1987) findings that transitivity is scalar, it should also be noted that there are different types of *in*transitivity as well. Unsurprisingly, the apparent neglect of intransitives as a construction class in the theoretical literature also extends to some of its narrower subtypes. One specific category of intransitives that the present dissertation is concerned with is the so-called zero-intransitive. The zero-intransitive is defined by its need for a "dummy", "pleonastic", or "expletive" subject such as *it* in *It is essential that you study for the final exam.* In German, it is the pronominal element *es* 'it' which acts as said expletive subject. In the traditional view, these types of subjects are semantically "empty" or "meaningless". It is claimed that non pro-drop languages simply require a dummy or expletive subject "even if it does not refer to any actual entity", in order to fill the syntactically required subject position (Velupillai 2012: 258). With respect to the perceived lack of semantics in traditional accounts, cognitive linguistic investigations take an alternative view and actually regard the so-called "dummy" pronoun as a meaningful element in the context of the overall construction. The often highly abstract or schematic semantic content of an expletive subject is a matter of construal, as discussed in the following.

In order to better apprehend the meaning of German *es* 'it' in weather expressions, we have to begin with Langacker's "conceptual archetypes" and the "setting/participant" distinction. In one such archetype, an expressed event or scenario is organized into settings and participants. Settings, in which events unfold, are typically rooms, buildings and/or geographical regions that do not participate in events, but rather host them. Locations are part of the setting

[19] The *Perfekt* tense of vi. is formed with *haben* 'have' while the *Perfekt* tense of vii.a. and vii.b. is formed with *sein* 'be'.

(points or areas) and therefore do not in themselves constitute any interaction with participants. Conversely, participants are typically people or discrete physical objects that interact with one another but not with locations; they merely occupy locations (cf. Langacker 2008: 11.1.1).

Propositions about meteorological phenomena like precipitation and temperature cannot be neatly divided into discrete entities like setting, process, and participants and, thus, the subject of such clauses is considered to have "non-participant status", which consequently qualifies such sentences as intransitives (Langacker 1991: 365). In such constructions, *it* (or *es* in German) is considered to have maximal generality. In making this point, Langacker cites Bolinger:

> [*it*] is a 'definite' nominal with almost the greatest possible generality of meaning, limited only in the sense that it is 'neuter' … It embraces weather, time, circumstance, whatever is obvious by the nature of reality or the implications of context. (Bolinger 1977: 84–85)

Langacker not only credits Bolinger with one of the more serious efforts at discerning the meaning of *it* but agrees with Bolinger's view in analyzing *it* as having maximal generality, rather than merely stipulating *it* as semantically meaningless. He offers one refinement, however, stating that "the 'ambient' sense of *it* designates an 'abstract setting' with 'non-participant status'" (Langacker 1991: 365). Ruwet (1991) discusses Bolinger's contribution of defining *it*, but does not seem convinced of Bolinger's arguments. Ruwet also discusses Darden's (1973) contributions but points out that Darden later abandoned his hypothesis in favour of Bolinger's (cf. Ruwet 1991: 97) which is why Darden is not discussed here. Based on Bolinger's statement that "any word which a language permits to survive must make its semantic contribution (1977: ix), Ruwet argues that Bolinger is inconsistent considering that he does not discuss the pro-drop cases, such as Italian *piove* 'it is raining' versus *it is raining* or *il pleut*, and thus the "missing" *it*. However, Ruwet's criticism would have more heft to it if that comparison could be made within a single language rather than between two languages. Ruwet wonders what Bolinger would make of the difference between *il pleut* 'it is raining' *and je (*le) regarde pleuvoir* 'I watch (*it) rain' and the semantic contribution of *it*. This is indeed another question this study is concerned with. Possible translations in German could be *Ich sehe es regnen* 'I see it raining' (inf.) or '*Ich sehe es regnet!* 'I see it is raining!' or *Ich sehe dem Regen zu* or 'I watch the rain (n.)' vs. *Ich sehe*

dem Regen beim Regnen zu 'I watch the rain at (the) raining' (*sehen* 'to see' vs. *zusehen* 'watch/look at'). Ruwet further states that Bolinger would have to find a difference in meaning between *il pleut* 'it rains' and *ça pleut* 'that rains', which seems to be the same as between the already mentioned German *es regnet* and *das regnet*. I cannot say whether Bolinger would agree, but the difference is quite obviously a regional one or one of register, both of which of naturally elude a purely structural analysis.

Langacker and Bolinger, of course, are not the only scholars who have concerned themselves with the meaning of *it*. Sullivan, using a construction grammar approach in her 2007 UC Berkeley dissertation, shows that the meaning of expletive subjects can be metaphorically mapped. One of the examples she gives is the following, where the cataphoric *it* is later succeeded by its conceptualization as *raingod* and *him* (cf. Sullivan 2007: 208–209, emphasis mine).

(24) *And it decided to rain. The rain god had apparently decided to keep us company. And we couldn't get rid of him the whole day.* (vinodgk.blogspot.com/2005_06_01_archive.html)

Apart from *it* being mapped onto some entity responsible for the rain, languages frequently employ agentives, such as deities or natural phenomena (like clouds or the sky), as the causers of MEs, some of which are discussed in Chapter 4, Section 4.3.1.

Lastly, studies like Smith (2002) demonstrate not only the meaningfulness of German *es* in all its uses, but its polysemy as well as its iconic motivation in a variety of constructions. Smith examines the uses of *es* 'it' in detail, relying on the framework of tenets and assumptions available in Cognitive Grammar (Langacker 1987, 1991). He presents a complex network of senses of *es* 'it' starting from the two categories of anaphoric and cataphoric *es* 'it' (Smith 2002: 103). He assumes the anaphoric third-person singular neuter *es* 'it' constitutes the prototype sense from which a variety of other senses can be semantically extended. These include, for example, "unspecified things"[20] and "reified processual relations".[21] On the cataphoric side we

---

20 *Er hat es gut/schlecht.*
he has it good/bad
'He's well/badly off.' (Smith 2002: 74)

21 [*In der Klasse ist gestohlen worden*,]$_{I}$ *aber keiner will es$_{i}$ getan haben*
in the class(room) is stolen become but no one wants it done have
'Something was stolen in the classroom, but no one admits to doing it [the process of stealing something].' (Smith 2002: 75)

find "concrete settings"[22] (spatial or temporal) and "abstract settings"[23] (which do not evoke the conceptualization of any actual physical setting). In Smith's categorization, weather expressions fall under concrete settings where *es* 'it' is a facilitative *es* that profiles restricted settings.[24] This means that speakers clearly link the meteorological event that is happening to a surrounding physical environment which facilitates or "hosts" the event. This is why such clauses do not necessarily require place adverbials and, moreover, can be analyzed as profiling a setting (cf. Smith 2002: 3.2.2).

In conclusion, the "expletive subject" that we find in so-called impersonal constructions including MEs, is not semantically empty, but is associated with the least maximal generality while facilitating the profiling of concrete restricted settings. The discussion of what *es* 'it' is profiling in MEs is important in order to motivate a semantic account of constructional meaning based on actual usage. Moreover, comparing the constructions in which precipitates and other weather words can occur, I investigate the respective semantic contributions of argument and predicate, in terms of their semantic "heaviness". *Es* 'it' clearly is a "light" argument, where we might expect the verb to do all of the heavy lifting semantically, as in *es regnet* 'it is raining'; whereas *Regen fällt* 'rain is falling' has a semantically "heavy" argument, while the verb is relatively light and multi-purpose. ARG-PRED-constructions, like *Regen regnet* 'rain is raining', contain two semantically heavy elements. The semantic notions of "light" and "heavy" are continuously used throughout the remainder of this dissertation. One of the goals of the thesis is to illustrate the complexity of "simple" intransitive clauses by examining one of their most prevalent exponents in the German language: meteorological events or weather constructions. The focus lies on the semantic contribution of the different components of the constructions (particularly the interaction of lexis and construction type), while simultaneously considering

---

22 *Es steht eine Vase auf dem Tisch.*
it stands a vase on the table
'There's a vase on the table.' (Smith 2002: 78)

23 *Es ist sicher, dass Hans gekommen ist.*
it is certain that Hans come is
'It's certain that Hans came.' (Smith 2002: 93, his example [34])

24 The second construction that falls under the category of "facilitative *es*" is the so-called reflexive impersonal passive construction:
Es tanzt sich gut hier
it dances-3SG REFL good here
'There is good dancing here.' (lit. 'It dances itself good/well here') (Smith 2002: 93, adapted from his example [26])

facets of the usage condition such as genre and speaker demographics, and extra-propositional content.

1.5 Fundamentals of cognitive linguistics and usage-based approaches to the study of meteorological expressions and why I chose a corpus-based approach

This section provides an overview of the basic tenets of the broad approach known as cognitive linguistics, which is a group of theories or approaches that share the same assumptions rather than a single theory. Cognitive linguistics encompasses most construction grammars and Langacker's Cognitive Grammar. While a lot of research in linguistics has been carried out following the idea that the system of language (grammar) should be examined separately from language usage (cf. de Saussure's *langue* vs. *parole* (2011 [1916]), Chomsky's *competence* vs. *performance* 1965), usage-based approaches reject this traditional distinction and the innateness hypothesis generally assumed by generative grammarians. Instead, they take language to be a dynamic, constantly changing system, where a speaker's grammar is continually restructured due to the demands of usage. Usage-based approaches can be undertaken without being tied to any specific grammatical framework, but it is particularly compatible with cognitive linguistics and construction grammar as usage-based theory is rooted in cognitive linguistics (Bybee 2010; Goldberg 2006; Langacker 1987); in fact, the term "usage-based" was introduced by Langacker himself (1987).

Langacker notably also developed Cognitive Grammar, or CG, which occupies a central niche in cognitive linguistics and is a comprehensive, elaborate theory of syntax and semantics (1987, 1991, 1999, 2000, 2008). His research on CG began in 1976 but can be said to have developed simultaneously with construction grammars (Langacker 2005: 157). One of the core principles of CG is that "language is grounded in language-independent cognitive processes such as association, automatization, schematization, and categorization" as well as embodied human experience (Hoffmann and Trousdale 2013: Introduction). In CG, the counterpart of what constructionist grammarians refer to as "constructions" (as illustrated below), Langacker calls "symbolic assemblies" (Langacker 2008), while "a construction is simply an assembly of symbolic

structures" (Langacker 2003: 50).[25] In CG, there is also greater emphasis on the cognitive principles and mechanisms that underlie the grammar and that are not specific to language, such as vision and analogy. These domain-general cognitive skills facilitate language acquisition, which is an ongoing process rather than pre-existent (see Tomasello 2003 and Goldberg 2006 for a detailed account of a usage-based theory of language acquisition). Two major claims of CG are that: "(1) general cognitive processes are fundamental to grammar; and (2) the emergence of grammar as a system of linguistic knowledge is grounded in language use" (Evans and Green 2006: 115).

Among the many concepts that Langacker detailed, there are the terms **"**profile", "base", and "domain", which he proposed specify relations in the theory of semantics in CG. The "domain" is a broader, more abstract conceptual setting whereas the "base" is a more immediate conceptual setting. Croft and Cruse (2004) liken the profile-base relation to a part-whole relation. The definition of ELBOW, for example, only makes sense against the concept of BODY PARTS. A domain is thus "a semantic structure that functions as the base for at least one concept profile" (2004: 16). The "profile" is the focus of attention: "a substructure that is elevated to a special level of prominence within the base" (Langacker 1991: 5) or the elements or features of the base that are foregrounded by the expression. According to Langacker, "the semantic value of an expression resides in neither the base nor the profile alone, but only in their combination: it derives from the designation of a specific entity identified and characterized by its position within a larger configuration" (1987: 183). Hence, the same base can provide different profiles. For example, the verbs in *it rained* and *it was raining* share the same base, but the former profiles a bounded event in the past whereas the latter profiles only a subsection or ongoing phase of the event. And while verbs profile the event, a noun, such as *rain*, reifies the process and profiles the ME as a 'thing'. Thus, while in the world of lexis, the profile of an expression is simply what the expression designates (Langacker 1990: 208–209), what is profiled in the world of grammar is a relationship. Both, relationships and things can be profiled. This highlighting of conceptual content is comparable to visual focus. Profiling is relative, though, and a matter of degree. Last but not least, profiling is not objective but depends on a speaker's proficiency and knowledge of conventionalization. It is no surprise, then,

[25] Langacker's "construction" thus has a broader meaning than that used in "explicit Construction Grammars". In CG, Langacker defines a construction as "either an expression (of any size), or else a schema abstracted from expressions to capture their commonality (at any level of specificity)" (2003: 44).

that speakers will choose one construction over another in a certain situation (e.g., *es regnet* 'it is raining' rather than *Regen fällt* 'rain is falling') presumably because they conceptualize an event differently (i.e., they invoke a different construal) and consequently choose to profile a different part of the base in this expression of the event.

Using alternate grammatical constructions allows us to verbalize the same phenomenon expressing "qualitatively different mental experiences" (Langacker 1987(I): 107) or what Langacker also terms "construal". It is crucial, in semantic analyses of both lexicon and grammar to take construal into account, as "the meaning of an expression is not determined in any unique or mechanical way from the nature of the objective situation it describes" (Langacker 1987(I): 107). In this dissertation, I will make use of concepts such as construal, the setting/participant distinction, as discussed in Section 1.3, and profiling, which I also refer to as foregrounding (and backgrounding where an event is not profiled, such as the MEs in prepositional constructions, cf. Chapter 4 of this dissertation). Alongside these concepts, I will also make use of the concepts of construction grammars (CxG), as described in the following paragraphs.

As cognitively grounded approaches, construction grammars do not only share many assumptions with one another, but also with other strands of cognitive linguistics. Accordingly, they maintain that "if two constructions are syntactically distinct, they must be semantically or pragmatically distinct" (Goldberg 1995: 67) and "grammar is assumed to be symbolic in principle; every distinction in form reflects a distinction in our conceptualization of the world" (Zelinsky-Wibbelt 1995: 45). Thus, Ruwet's examples from the last section (*the day is dawning* vs. [*the*] *dawn is breaking*), too, must have different meanings. One difference, for example, lies in pragmatics as the discriminating factor, which according to Goldberg is tied to "particulars of information structure, including topic and focus" (1995: 67). Construction grammar (CxG), and the Frame Semantics approach to linguistic analysis have their historical roots in Berkeley (Berkeley Construction Grammar, or BCG) and particularly in the work of Charles Fillmore. The foundations of CxG were described and developed in detail by Goldberg in her seminal work of 1995. I already mentioned the *FrameNet*, an online lexical database of English based on the theory of Frame Semantics above (in the summary of Liégeois's research). It will come into play in Chapter 3, where many of the modifiers that the analyzed data were coded for (TIME, LOCATION, QUANTITY, MANNER, etc.) correspond to the core and non-core elements established by *FrameNet* for frames such as "precipitation" (cf. fn.9).

In actuality, a whole family of construction grammar frameworks have emerged from the origins of BCG and CG. Most of them share fundamental assumptions (adopted by the study at hand), the core one being the following: the basic unit of linguistic organization is a construction.[26] Originally, constructions were considered form-function/meaning pairings that are non-compositional in some aspect, or "at least partially arbitrary" (Croft 2001: 18). Later, this definition was extended to include any cognitively routinized or conventionalized form-meaning patterns: ''[F]acts about the actual use of linguistic expressions such as frequencies and individual patterns that are fully compositional are recorded alongside more traditional linguistic generalizations'' (Goldberg 2006: 45). There are no underlying transformational or derivational components involved in what are monostratal representations and no phonologically empty elements posited. Form and meaning are linked via symbolic correspondences (see Figure 2).

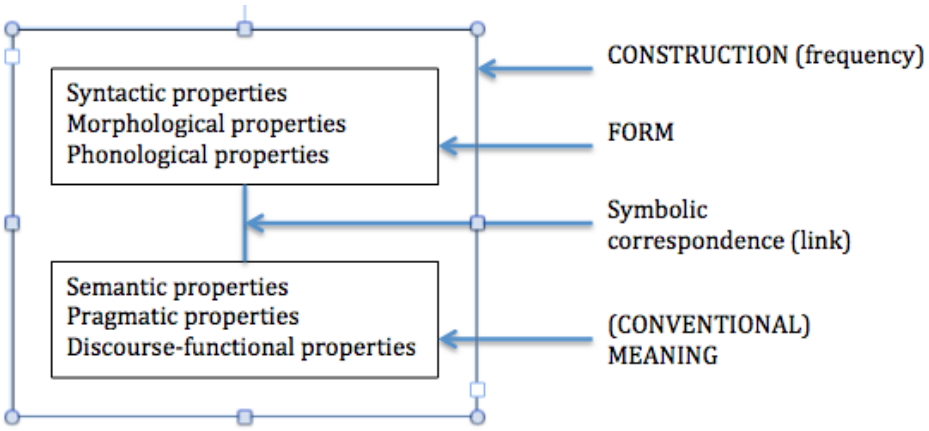

Figure 2. *The symbolic structure of a construction (Croft 2001: 18)*

Thus, instead of following the traditional view that lexicon and syntax are distinct linguistic categories, constructions are part of a lexicon-syntax continuum. All dimensions of language (i.e., syntax, semantics, pragmatics, discourse, morphology, phonology, prosody) are perceived to contribute equally in shaping the form and meaning (and use) of linguistic expressions. Words

[26] For example, *Radical Construction Grammar* (RCG) (Croft 2001), designed for typological purposes and allowing for cross-linguistic factors and Goldberg's *Cognitive Construction Grammar* (CCG) (1995, 2006; Boas 2013) "which aims at providing a psychologically plausible account of language by investigating the general cognitive principles that serve to structure the network of language-specific constructions" (Boas 2013: 233). Other construction grammars are *Sign-Based Construction Grammar* (SBCG) (see, e.g., Michaelis 2013), which is a blend of *Berkeley Construction Grammar* and *Head-Driven Phrase Structure Grammar* (HPSG) (Pollard and Sag 1987, 1994), *Embodied Construction Grammar* (ECG) (Bergen and Chang, 2005), aimed to support embodied models of language use and learning. For an overview see Hoffmann and Trousdale (2013).

and clauses both have constructional properties and therefore form constructions. Correspondingly, construction grammars represent a holistic approach that deems all linguistic expressions equal and semi-independent due to particular usage conventions rather than universalist phrase structure requirements. More unusual "peripheral" idiosyncratic structures particular to a language (e.g., idioms) are understood to be acquired the same way as more general, "basic" or core ones (e.g., transitives or declaratives) (cf. Fried and Östman 2004). Knowing a language means knowing a language's structured inventory of constructions, sometimes called its **"**constructicon".

Constructions differ in size and complexity. On the lexicon-grammar continuum, constructions differ in terms of their schematicity. They range from very specified, i.e., "fully filled and fixed", to "fully schematic". "Fully filled and fixed" can be exemplified by a lexical item like *oxen* which allows for no variation. A "fully schematic" construction can be exemplified by the regular past tense, $stem_V$-PAST (*work-ed*, *paint-ed*), corresponding to syntactic rules in other theories. Along the continuum there are also constructions that are "fully filled and partially flexible" (e.g., *go*[tense] *ballistic*), and "partially filled" (e.g., *the* [AdjP] – *the poor/beautiful*) (cf. Fried 2015). Following the continuum from the fully concrete to the highly schematic, constructions are hierarchically organized in inheritance networks that overlap and complement one another. They are related by shared properties. Inheritance can be partial or complete. Last but not least, another crucial commonality of construction grammars is that they are usage-based. As has already been alluded to, this means that constructions, or language structures, emerge out of language use. This is true for individual words, but it also applies to more complex levels of grammar, such as multi-unit constructions.

The most frequent constructional patterns are entrenched**.** The notion of "entrenchment", too, goes back to Langacker (1987: 59) and refers to the extent to which a cognitive structure is routinized and automated. The more speakers use a structure, the more it becomes entrenched while disuse of a structure has a negative impact.[27] Entrenchment is furthermore linked to conventionalization which is defined as "the continuous mutual coordination and matching of communicative knowledge and practices, subject to the

[27] See Schmid (2015) for a general outline of the theoretical model of entrenchment and conventionalization, where the former differs slightly from Langacker's understanding of entrenchment, the sociopragmatic processes underlying entrenchment being association, routinization, and schematization.

exigencies of the entrenchment processes taking place in individual minds" (Schmid 2015: 10), or also long-term language change. Both entrenchment and conventionalization are affected and linked by usage and in turn influence usage, where frequency of usage is of utmost importance (Schmid 2015).

Frequency plays a fundamental role in the usage-based approach, as it both shapes the linguistic system and is also a result by way of a feedback loop. Highly frequent items or patterns, i.e., items that have a high turnover of repetitions, are more entrenched. According to Langacker (1987), entrenchment in cognitive organization is scalar.

> Every use of a structure has a positive impact on its degree of entrenchment, whereas extended periods of disuse have a negative impact. With repeated use, a novel structure becomes progressively entrenched, to the point of becoming a unit; moreover, units are variably entrenched depending on the frequency of their occurrence.
>
> (Langacker 1987: 59)

Repetition not only plays a key role with regard to entrenchment, but also with regard to conventionalization as it expedites diffusion across different communicative situations, such as written and spoken discourse, different speakers, and different registers. According to Schmid, diffusion (a key process of conventionalization) starts at the point where a new utterance/construction is used for the first time in a new communicative situation. Moreover, it needs to be repeated in different communicative situations for conventionalization to take hold (2015: 18). That is, ideally the innovative construction is read the first time (rather than heard) and then repeated in either an oral or written situation. In terms of frequency, it is not just the number of speakers who repeat the innovation that plays a key role, but that the innovation is furthermore repeated across a number of different contextual situations, such as social contexts, genres, and registers. Accordingly, we will see that the predicate *schneeregnen* 'to sleet' is, in fact, not (yet) sanctioned by the *Duden* and does not readily enter the "default" predicate type construction [*es*+V_{PREC}] in standard German (cf. Chapter 3 of this dissertation). It is, however, used informally not only by speakers online, as in (25), but also in more formal print media, like the example in (26) from the *Bayerische Staatszeitung* 'Bavarian National Newspaper', to give just two examples.

(25) *Igitt:* ***es #schneeregnet*** *- dachte es soll #Sommer werden?!*[28]
'Yuck: **it is #sleeting** – (I) thought it was supposed to be #summer?!'

(26) ***Es schneeregnet****, es ist kalt, so richtig ungemütlich an diesem Montagmorgen (...).*[29]
'**It is sleeting**, it is cold, just really unpleasant this Monday morning (…).'

In more informal settings, speakers even go one step further and reduce the compound verb *schneeregnen* to the blend *schneegnen* (with apparently the same meaning) as documented in examples (27) and (28). The former was found on the social media platform *Instagram* where the author of the post put the word *SchneeRegen* [sic] 'SnowRain' in parenthesis after *schneegnet*, presumably to clarify the meaning of the neologism.

(27) *Es ist trüb* ***es schneegnet*** *(****SchneeRegen****) es ist kalt aber nicht kalt genug...*[30]
'It is dull **it is sleeting** (**snow rain**) it is cold but not cold enough.'

In the example in (28), from an online message board, the user points out that their use of *schneegnet* is a neologism and not a typographical error (after another user earlier on in the conversation used *schneeregnen* as a verb). Both inflected forms *schneeregnet* and *schneegnet* have their own hashtag on *Twitter*. A hashtag search on *Twitter* returns five items for *schneeregnet* (with no hashtag for the non-finite form) and almost twice as many, i.e., nine, for the neologism *schneegnet* (plus two non-finite instances of #*schneegnen*). Additionally, a quick *Google* search returns 2,200 results for *es schneeregnet* and 122 for *es schneegnet*.[31]

(28) *oberfranken: +1°,* ***es schneegnet*** */kein tippfehler, sondern neologismus!*[32]
'Upper Franconia: +1°, **it is sleeting**. /no typo, but neologism! '

The returns from the corpora as well as the examples in (25)–(28) are far from the case of an isolated *hapax legomenon*. As speakers come across more instances of the predicate

[28] Source: https://twitter.com/versustommy/status/695282663586951169 (last accessed July 4, 2019).
[29] Source: *BSZ,* www.bayerische-staatszeitung.de/staatszeitung/kommunales/detailansicht-kommunales/artikel/begehung-mit-schwachstellenanalyse.html#topPosition (last accessed July 4, 2019).
[30] Source: www.instagram.com/p/BtYtpa7DU0k/ (last accessed July 4, 2019).
[31] The search was conducted on the German *Google* domain for the exact phrase/construction by putting it in quotes, which will yield only pages with the same words in the same order as what is in the quotes. The search was conducted on July 1, 2019.
[32] Source: https://forum.mods.de/bb/thread.php?TID=193996&page=5 (last accessed July 4, 2019).

construction with *schneeregnen* 'to sleet' filling the verb slot of the predicate construction—an "ongoing production-perception feedback loop" (Bybee and Beckner 2009: 833)—they may adopt it into their inventory of conventionalized constructions enough so that in due course it might be added to the *Duden* and consequently become part of the standard grammar of German. For the time being, however, neither *es schneeregnet* nor *es schneegnet* are prescriptively acceptable (although both are acceptable to the author). Instead, a standard speaker must use an argument construction as in *Schneeregen fällt* 'sleet falls' or *es gibt Schneeregen* 'it gives sleet' (cf. Table 1), which points to an imbalance in the use of construction types and their interaction with different precipitate types.

The uneven distribution of ME constructions (cf. Table 1) is furthermore evident in other usage events. Consider (29) and (30) below (taken from the *deTenTen13*). In (29), the speaker, describing a typical November day (in Germany arguably the dreariest month of the year weatherwise) chooses an argument construction for the fourth predication, *es fällt Nieselregen* 'it falls drizzle', rather than continuing the preceding list of adjectival predicate type constructions with, for example, *es ist regnerisch* 'it is rainy', or even *it is nieselig* 'it is drizzly' (although *nieselig*, like *schneeregnen* 'to sleet' is not sanctioned by the *Duden* either; nevertheless, *nieselig* is part of online dictionaries and speaker usage: *Es ist nieselig* 'It is drizzly [weather]'). They could, nevertheless, simply have chosen the "default" expletive predicate subtype *es nieselt* 'it is drizzling'.

(29) *Es ist kalt. Es ist feucht. Es ist dunkel. Nieselregen fällt. (…) November eben.*
'It is cold. It is wet. It is dark. Drizzle (lit. 'drizzlerain') is falling. Simply November.'

This is particularly interesting as the speaker in (30) describes the month of November in a similar manner where they, indeed, use the adjectives *regnerisch* 'rainy' and *neblig* 'foggy', but conclude the portrayal with the argument construction *und gibt's Schnee* 'and there is snow', lit. 'and gives it snow'. Take also note how the first predicate is inflected for future tense (*sein* 'to be' *werden* 'will be') while the predicate of the second main clause is inflected for present tense (*es gibt* or *gibt's* 'it gives' instead of *wird es geben* 'it will give'). This may be an indicator of the use of the [*es+geben*] construction with future or imminent events. Consider also utterances such as *Es gibt Essen!* lit. 'it gives food', which is readily used as a call to dinner ('Dinner's ready!') or *Das gibt gleich was!* lit. 'That gives something in a moment!' while looking at dark clouds indicating that rain is imminent.

(30) *Im November hingegen **wird es regnerisch, neblig** und **gibt's Schnee***.
'But in November it **will be rainy**, **foggy** and **it gives snow**.'

Since German does not have an adjectival *Schnee* 'snow' equivalent for *regnerisch* 'rainy', the use of a parallel adjectival construction would be infelicitous in (30). The German language has the adjectives *verschneit* und *schneeig*, both meaning *von Schnee bedeckt* 'covered with snow' rather than *snowy* as in 'when snow is falling'.[33] There are many such examples in the data, where the speaker lists different MEs by employing different construction types for different precipitate types. Once again, not all construction types are equally available for all precipitate types, although we would expect it to be this way. That is, there should be as many predicate constructions appearing with *snow*, and *hail*, and *sleet*, etc. as with *rain* (or at least proportionally regarding the frequency of their natural occurrence). In the utterance in (31) from a private travel and hiking blog, the speaker chose parallel constructions by listing the ME-predicates without repeating the pronoun *es* 'it', which serves as the single subject for each predicate.

(31) ***Es regnete**, **schneite**, **hagelte** und **graupelte** im steten Wechsel*.[34]
'**It rained**, **snowed**, **hailed** and **graupled**/sleeted in constant succession.'

[33] These are the definitions according to the *Duden*. According to personal communication with several German native speakers, *verschneit* could mean both things as in *Es ist verschneit*/*Die Stadt ist verschneit* 'It/The city is covered with snow' but also *verschneit* as a modifier as in *Es ist ein verschneiter Tag* 'It is a snowy day', meaning a day where snow is falling, even if it is not snowing the whole day. According to learning platform *Quizlet.com* (https://quizlet.com/10456860/weather-flash-cards/, last accessed October 12, 2019), the translation for 'It is snowy' is given as *Es ist schneeig* but provided with the additional information "*im Deutschen nicht üblich*" 'not used in German'. It is not clear to me whether it is meant that *schneeig* is not used in standard German but possibly in Swiss German or other dialects. It is also odd to offer learners a translation when it is claimed that the word is not used. The German word *verschneit* is not given. The online dictionary *dict.cc* translates *schneeig* as 'snowy' or 'niveous' and *dict.leo.org* translates *schneeig* as *'snowy'* (both last accessed October 12, 2019). The odd example can, of course, be found: *Der Frühling ist hier ziemlich kühl und der Sommer ist nass, warm (…), regnerischer Herbst, der Winter ist nicht zu kalt, **es ist schneeig**, mit Tauwetter*. 'Spring is fairly cool and summer is wet, warm (…), rainy autumn, winter is not too cold, **it is snowy**, with thaw-weather.' (Source: https://de.fehrplay.com/obrazovanie/83640-prirodnaya-zona-ukrainy-step-lesostep-smeshannye-lesa-gory.html [last accessed May 30, 2019).

[34] Source: https://wanderlustig2019.wordpress.com/2021/03/26/wintersonne-in-aschau-wintersun-in-aschau/ (last accessed May 20, 2024).

An intentional choice—against assumed "default" constructions—may be founded in conscious style choices, which may be more pronounced in LITERARY texts, or different construals. Keep in mind, that there does not seem to be a level playing field between precipitates to begin with as attested by the lack of the prescriptively sanctioned verb *schneeregnen* 'to sleet' and, to some extent, also *hageln* 'to hail', which is predominantly used in a figurative manner (cf. Sections 3.4.3 and 4.4, as well as Andrason [2019] and Andrason and Visser's [2019] findings for Polish and isiXhosa) .

The examples in (25)–(31) are authentic utterances, which were not created for the purpose of the dissertation at hand. As Stefanowitsch notes "there is no way to study usage other than by looking at usage data" (2011: 285). Connected to usage-based approaches is the importance of authentic data obtained by means of a corpus or corpora, a premise which is explored in the remainder of this subsection with due regard to how cognitive linguistic approaches and corpus linguistic methodologies complement each other.

It follows from the inextricable linkage of the linguistic system with usage that any linguistic analysis and theory of language should be based on recorded instances of actual language use rather than on introspection, made-up or constructed examples, examples from standardized grammars, or acceptability judgments of what may or may not be (theoretically) possible. While intuition and introspection are both interesting and a necessary precursor to any linguistic investigation, language users (including linguists) are often oblivious to the existence of certain items and structures, let alone their frequency or collocations (cf. Bybee 2006; see Wasow and Arnold 2005 for reliability of judgment data by linguists). Introspection does not give "evidence about usage" (Sinclair 1991: 39). However, intuition and introspection also have their place when it comes to analyzing and interpreting the data.

> [...] artificial data can have a place in modern corpus linguistics. Yet it should always be used with naturally occurring data which can act as a control, a yardstick if you will. Corpus linguistics is, and should be, a synthesis of introspective and observational techniques, relying on a mix of artificial and natural observation. (McEnery and Wilson 2001: 16)

Both cognitive and constructionist approaches furthermore demand an analysis that, in the Firthian tradition, aims for understanding a lexical item or pattern in its context of use and not in isolation. From the morphological inflection of a lexical item to its collocations (i.e., any other lexical elements that co-occur with it in a phrase) to full-blown syntactic structures, all of

these factors contribute to the overall function or meaning expressed by the linguistic item, be it natural or constructed. Thus, examining the entire lexico-syntactic frame in which a construction occurs in authentic manifestations naturally leads to choosing a corpus-based approach. The corpora which were chosen to serve as the data source for this dissertation, i.e., the *deTenTen13* and the *DWDS* corpora, are described in Chapter 2. Meanwhile, let us consider how to reconcile corpus linguistics and cognitive linguistics.

In a well-known article, Fillmore (1992) sketches the caricature of an "armchair linguist" and a corpus linguist. In this caricature, the armchair linguist does the thinking while the corpus linguist merely crunches numbers on facts. They do not communicate, but when they do, they are suspicious of or simply not interested in the other's work. While this may have been true to some extent in the past, more and more cognitive linguists make use of corpora (see, e.g., Gries and Stefanowitsch 2006; Stefanowitsch and Gries 2006; Gries 2012, 2015; Luodonpää-Manni, Penttilä and Viimaranta 2017) and quantitative methods in general, to a degree where Janda announces "the quantitative turn" in linguistics (cf. Janda 2013).[35] In fact, the use of corpus data may be one of the most routinely used methodological approaches in cognitive linguistics and CxG.

Corpus linguistics and cognitive linguistics not only harmonize, but they do in fact complement each other (cf., e.g., Gries 2012; Guilquin 2010). In a position paper, Gries (2012) lays out the many ways in which both fields are compatible and how corpus linguistics is immediately relevant to both cognitive linguistics and psycholinguistics.[36] He argues how both type and token frequencies are of interest to cognitive linguists and psycholinguists. To the former, token frequencies are of interest because they correlate with "entrenchment (…), phonetic reduction and development of new forms" while type frequencies "correlate with (morphological) productivity and language change" (2012: 49). To the latter, token frequencies are of interest because they "correlate with ease/earliness of acquisition (…), lexical decision tasks, word naming and picture naming" while type frequencies correlate with "the productivity of constructions in first and second language learning" (Gries 2012: 49). Other commonalities that Gries mentions include—but are not limited to—advocating for a lexicon-syntax continuum and a shared interest in dispersion (which has implications for language acquisition and

[35] Even proponents of usage-based models get by without using naturally occurring data, cf. Lakoff and Johnson's *Metaphors We Live By* (1980) and Langacker's *Foundations of Cognitive Grammar* (1987, 1991).
[36] Gries, of course, understands corpus linguistics as a methodology rather than a theory (cf. Gries 2012).

psycholinguistic experiments), the similarities between Sinclair's (1991) *Idiom Principle* and Langacker's "unit" (1987), as well as the resemblance of Hunston and Francis's (2000) patterns with Goldberg's (2006) constructions (cf. Gries 2012: 49–51).[37]

Regarding the strategies employed by corpus and cognitive linguistics, Guilquin (2010: 15) points out that while it "is true that corpus linguistics extracts categories from the data and cognitive linguistics posits them beforehand (…), this is no different from saying that corpus linguistics starts with the data while cognitive linguistics starts with the theory". The crucial commonality, however, is that both frameworks recognize the existence of categories and, more importantly, acknowledge the non-discreteness of categories. This overlap is evident in the above-mentioned lexicon-grammar continuum. Cognitive considerations afford the "number crunchers" a theoretical sophistication that is often claimed to be wanting in corpus linguistics (Guilquin 2010: 16). In turn, the use of corpora in cognitive linguistics validates its claim of being "usage-based", and the evidence obtained from corpora substantiates cognitive theories, postulations and generalizations made (cf. Schönefeld 1999).

Among others, proof for these claims can be found in studies on the dative alternation, which has been the target of much syntactic analysis (if not obsession) for decades. A speaker's choice between two or more constructions is based on a number of factors for which psycholinguistics and corpus studies have both accrued solid evidence. Factors that influence the choice are, for example, verbal semantics (Levin 1993; Gries 2005; Bresnan, Cueni, Nikitina, and Baayen 2007), animacy of the recipient (e.g., Bresnan and Nikitina 2009), the cognitive load (in terms of complexity) of the theme/recipient (cf. de Cuypere, de Coster, and Baten 2014), and information structure concerning the theme/recipient, that is, is the referent new or given information (e.g., Bresnan et al. 2007). Other studies that demonstrate the reliability and usefulness of corpus-based methods for a usage-based quantitative approach, as adopted in this dissertation for the examination of the lexico-syntactic behavior of MEs, are those on "basic" verbs, such as Newman and Rice (2006) on EAT and DRINK, or English posture verbs SIT, STAND, and LIE (2001, 2004), etc. In studying the (in)transitivity of the verbs EAT and DRINK, based on thousands of instances from both the written and spoken *British National Corpus* (BNC),

[37] "A language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments" (Sinclair 1991: 110) vs. "a structure that a speaker has mastered quite thoroughly, to the extent that he can employ it in largely automatic fashion, without having to focus his attention specifically on its individual parts for their arrangement […] he has no need to reflect on how to put it together" (Langacker 1987: 57).

Newman and Rice (2006) found that these two diathesis-alternating verbs behave differently not only in "their argument structure and in the interpretation of their subject and object depending on genre, register, or modality (by which we mean spoken vs. written corpus)" but more importantly that the following features are closely connected to the verb's valency: TAM (tense/aspect/mode) markers, "person, number, specificity of the subject, and the semantic properties of their propositional and extra-propositional collocates" (2006: 226). The actual usage of items reveals that "inflected verb forms have their own semantic and constructional properties" (2006: 255). While weather events are usually restricted to the third person inflection, Newman and Rice's (2006) findings, nevertheless, translate to MEs in that the linguistic features cited in the paragraph above are also strongly tied to the variety of ME constructions available as well as the different precipitates that enter the respective construction. Each ME construction comes with a cohort of meaning extensions and usages and, most importantly, specific constructional elements that may be both propositional and extra-propositional. While events share certain contextual features, we can see that the manipulation of TAM and semantic features reveal sharp selectional distinctions (cf. Chapter 3 of this dissertation).

This section has laid out the fundamental assumptions, concepts, and terminology of cognitive linguistics in general and CG, in particular, and construction grammars and how these frameworks operate within a usage-based approach. Furthermore, I shed light on the interrelatedness of cognitive linguistics and corpus linguistics and the importance of using authentic data when operating within a usage-based approach. The next section, which concludes the introductory chapter, comes back full circle to Talmy's typology of motion events and involves the "last" linguistic approach relevant to the study at hand: lexical typology.

1.6 The emerging field of lexical typology

The field of lexical typology, while not yet clearly defined, is a growing subdiscipline that focuses on crosslinguistic research of the lexicon and studies the organization of semantic fields (cf. Koptjevskaja-Tamm, Vanhove, and Koch 2007 and Rakhilina and Reznikova 2014). One generally agreed upon understanding of lexical typology, as advocated, for example, by Koptjevskaja-Tamm and others (cf. Koptjevskaja-Tamm 2008; Koptjevskaja-Tamm, Rakhilina and

Vanhove 2016; Rakhilina and Reznikova 2014), is that it deals with "characteristic ways in which language […] packages semantic material into words" (Lehrer 1992: 249). Lehrer, an early advocate of semantic typology, raised questions as to whether style, connotation, and semantic prosody should be part of the meaning of a term (cf. Lehrer 1974), which is the same view as advocated by construction grammars.

Anyone who has ever learned more than just the very basics of a new language will know that studying the lexicon and mastering the correct use of a lexical item is not a simple task. Translation equivalents for concrete items like *table*, *dog*, or *knife* are readily made, but simple translation falls apart quickly, as I hope to have demonstrated in Section 1.2, in the discussion about differences in motion-event packaging across the world's languages. More abstract words come with specific nuances, selectional restrictions, and semantic prosodies and associations that are harder to discern and master. In the domain of AQUA-MOTION, for example, Lander, Maisak and Rakhilina (2012) identify the sub-domains of FLOATING, DRIFTING, and SAILING as the most basic across languages. However, while English has a relative richness of AQUA-MOTION verbs (*float, drift*, *sail*, and *swim*; Indonesian has even more with fourteen AQUA-MOTION verbs), there are languages like Turkish, where there is an undeniable homogeneity, or paucity, of AQUA-MOTION verbs, having only one such verb (cf. Lander et al. 2012). Following Talmy, scholars working on AQUA-MOTION verbs distinguish between different semantic components, such as FIGURE, GROUND, MANNER, and PATH, but focus on the MANNER component. The systematic cross-linguistic variation found in the AQUA-MOTION domain (Koptjevskaja-Tamm, Divjak and Rakhilina 2010) does not support the claim that S-languages have a much richer MANNER lexicon than V-framed languages (Slobin 2003). Koptjevskaja-Tamm et al. (2010) find that Slavic languages (S-framed) have a much poorer AQUA-MOTION MANNER lexicon than presumedly "MANNER-poor" Romance V-languages. For example, Russian and Polish each only have two AQUA-MOTION verbs (*plyt'/plavat'* 'swim/float/sail' in Russian and *płynąć/pływać* 'swim/float/sail' in Polish). Observations of non-semantic equivalence like this have led to the emerging field of lexical typology, which attempts to bridge the fields of lexicography and grammatical typology. Rakhilina and Reznikova (2014) maintain that grammatical and lexical typology share the same attitudes and goals. Traditionally, grammar and the lexicon are considered to be dichotomous, with the former assumed to be structured and systematic (if not universal) and the latter a repository of language-specific idiosyncrasies. The lexicon is often regarded as too complex and unsystematic to afford much in the way of systematicity. Linguistic typology, on the other hand,

is traditionally based on the systematicity of grammatical phenomena (such as word order) as it is concerned with functions and features that vary regularly in a wide cross-section of languages. It is this systematicity in grammatical structure that presumably allows for overall generalizations and implicational hierarchies. However, there is systematicity not only in grammatical structures per se and the grammatical structure of the lexicon (e.g., word classes), but also in how the lexicon carves up domains of experience, such as the human body, motion, perception, and so on. Lexico-typological research thus tries to marry lexicography and typology by carrying out systematic research on the lexicon, asking what parameters underlie specific phenomena and domain categorization.

Talmy's seminal work on the typology of motion event encoding is considered one of the first and now classical typological studies of the lexicon. His typology, outlined in Section 1.2, has set the groundwork for the field of lexical typology by illustrating how different languages lexicalize meaning-form relationships in the general area of event realization. Talmy revealed the intricate interplay between various semantic elements and the surface elements with which they may be expressed and detailed their conflation patterns. Considering a sentence such as *Regen stürzt vom Himmel* 'rain is gushing from the sky' (*stürzen* 'plummet/plunge/fall/rush'), the verb is semantically complex and conflates the fact of MOTION, the MANNER of motion, and also the PATH of motion. If we contrast this with the semantically "lighter" *Regen fällt vom Himmel* 'rain is falling from the sky', the complex conflation of the fact of MOTION and the PATH remains. In comparison to the relatively violent *stürzen*, however, *fallen* is neutral and offers no special indication of the MANNER of motion, only direction. Similarly, *es regnet* 'it is raining' can be broken down into the conflation of the FIGURE of rain with the fact of MOTION while in *es schüttet* 'it is pouring', the fact of MOTION and MANNER are conflated, but the FIGURE (of rain) is implied at best, again based on our world knowledge that it must be rain that is pouring down.

Talmy's original survey demonstrated that there are systematic yet non-isomorphic relations between meaning and surface forms. He has been able to isolate the meaningful elements or components that constitute the makeup of lexical and grammatical items in motion events. Traditional typologies are commonly based on syntactic properties such as word order. The examples above (*es regnet* 'it is raining' vs. *es schüttet* 'it is pouring') are both "predicate types" of the "expletive subtype", if we follow by Eriksen et al.'s (2010, 2012) typology. However, following a Talmyan classification, we find diverging conflation patterns of the semantic elements underlying the predicates. Lexical items do not just carry ever-changeable

semantic meaning, but rather there is also a "grammatical structure" to the lexicon, which, according to Lehmann (1990) is:

> the least manipulable in a linguistic system (and the) least exposed to the discretion of the speaker**.** To be precise, we are led to the grammatical structure of the lexicon; not, of course, to the semantics of individual lexical items, their configurations in lexical fields or individual processes of word-formation. These are relatively changeable and short-lived, as has been recognized from the beginning of studies in lexical semantics. What is, however, fairly deeply entrenched in linguistic structure and resistant to change is the grammatical categories, features and relations which are coded in the lexicon and its classes. (Lehmann 1990: 165)

Thus, the way a language chooses to lexicalize concepts and specific categories of concepts directly "influences the shape that its grammar takes, because grammatical categories are essentially fed by lexical categories" (1990: 180).

As pointed out in Section 1.5, it is argued that constructions themselves carry meaning, a central thesis being that constructions are form-meaning pairings that exist "independently of particular verbs" (Goldberg 1995: 1). Let us consider the German ME construction *es gibt Regen* 'it gives rain', for example. In general, the German verb *geben* 'give' functions in a similar way as its English counterpart, *give*, and acts as the premier verb of transfer. However, *geben* also leads a more mysterious second life due to the fact that it has undergone extensive grammaticalization. In combination with *es* 'it', the third person singular pleonastic pronoun, *es gibt* has an existential reading, as illustrated in (32), with the accusative object (underscored) being interpreted as the subject (see Newman [1996, 1998, 1998a] for a more elaborate discussion of the linguistics of *giving* in general and German *es gibt* and its origin in particular).

(32) *Weltweit gibt es rund <u>zehntausend Zoos</u> (…).*
'Worldwide there are roughly ten thousand zoos (…).'

When *es gibt* is used in a weather expression, however, the overall construction can take on a future reading as already demonstrated in (30) (cf. also fn.33), especially where it is combined with adverbs such as *morgen* 'tomorrow'. *Google Translate*, similarly, translates *das gibt ein Gewitter* as 'there will be a thunderstorm' and not 'there is a thunderstorm'. Chapter 3 will show, though, that across the three major constructions analyzed (ARG+*fallen*, ARG+*geben*, *es*+PRED) and the four precipitate types, the [ARG+*geben*] construction occurs equally

frequently for past tense events (cf. Section 3.5). Whereas the existential (or habitual) meaning in English can be translated with 'there is', there is no future reading available in English. For the former, the utterance *There is a lot of rain in Seattle* acknowledges that Seattle averages more rainfall than other U.S. cities, but the same sentence cannot be used to forecast rain. In such cases, future tense marking is needed: *There will be rain today* or *There will be rain later today and tomorrow*.

These examples show how the interaction between lexicon and grammar leads to very different outcomes in terms of the form, meaning, and acceptability of MEs when they are examined as an organic whole rather than independently (cf. also Newman and Rice's findings mentioned above). Fortunately, lexical typological approaches have been afforded much more attention recently. The field of lexical typology, however, is still in its infancy and there does not seem to be consensus on or an exact definition of what lexical typology entails. In recent years, there has been an increased interest in lexical typology as demonstrated by research groups and research programs such as the "Categories across language and cognition" project at the Max Planck Institute for Psycholinguistics in Nijmegen,[38] Evans's chapter on semantic typology in *The Oxford Handbook of Linguistic Typology* (2010), or numerous publications centered around basic verbs of specific domains: a 2015 volume on TEMPERATURE expressions edited by Koptjevskaja-Tamm; Newman's edited volumes on GIVE (1998), posture verbs SIT, STAND, and LIE (2002), as well as verbs of consumption EAT and DRINK (2009); Maisak and Rakhilina (2007), Koptjevskaja-Tamm, Divjak and Rakhilina (2010) and Lander et al. (2012) on AQUA-MOTION; Evans and Wilkins (2000), and Viberg (2001) on perception verbs SEE and HEAR; Majid and Bowerman (2007) on CUTTING and BREAKING events and Majid, Enfield and van Staden's (2006) edited volume on BODY PARTS.

Studies in lexical typology tend to focus on different aspects concerning the semantics and structure of the lexicon, and address different questions, as discussed below. Talmy's research addresses the syntagmatic axis, which is concerned with the packaging of conceptual material into sequences of lexical items, as opposed to the paradigmatic axis, which is more concerned with packaging conceptual material into single lexical units (cf. Koch 2001: 1169–1171). Wälchli and Cysouw (2012) explore theoretical and methodological aspects of the lexical-typological approach (like Talmy, their work is concerned with motion verbs). Lehmann (1990) and Koch (2001, 2008) focus on lexicon-grammar interaction, while Juvonen and Koptjevskaja-

[38] www.mpi.nl/departments/other-research/research-projects/categories

Tamm (2016) deal with polysemy and semantic shifts. No matter where the main interest lies (in the grammatical structure of the lexicon or the semantics of the words), compared to typological research on grammar, phonetics, and phonology, lexical typology is woefully understudied.

One strong advocate of semantically based research in lexical typology is Maria Koptjevskaja-Tamm. She suggests an approach to lexical typology that takes three major perspectives into account. These are the *onomasiological perspective*, the *semasiological perspective*, and *lexicon-grammar interaction*. The onomasiological perspective and the semasiological perspective are basically two sides of the same coin, or complementary but inverse approaches to uncover the connection between meaning and words/expressions where relations are "not 1:1 (…), but n:n" (cf. Lehmann, n.d.).[39] Onomasiology considers concepts or domains first and examines how these are given lexical realization. As we have seen at the beginning of this section, regarding the domain of AQUA-MOTION, Turkish only has one verb to convey motion through water, while English has four verbs. For the domain of WEATHER, we can ask how many words are there to denote precipitation or a certain type of precipitation? As we saw earlier, Andrason (2019) points out that Polish has no equivalent for 'sleet' and needs to resort to paraphrases involving 'rain' and 'snow'. Semasiology, on the other hand, starts from the individual lexical item and looks at which concepts or meanings can be referred to by that form. If we look at *hageln* 'to hail', we find uses such as those in (33) and (34); the former literal and the latter figurative. In (33) *hagelte* 'hailed' refers to precipitation in pellets of ice. This ME meaning of '*it hailed'* would remain, even if the complement (*drei Zentimeter große Eiskörner*) was removed. In (34) though, we cannot remove the complement *Kritik*, as this would lead to a cognitive disconnect between the two sentences.

[39] www.christianlehmann.eu/termini/onomasiology&semasiology.html (last accessed April 14, 2017).

(33) *In Delmenhorst **hagelte es** drei Zentimeter große **Eiskörner** vom Himmel*.[40]
'In Delmenhorst, **it hailed** three centimeter big **hailstones** from the sky.'

(34) *Mit 63 Jahren hat eine Frau in Australierin* [sic] *ihr erstes Kind zur Welt gebracht. In Australien **hagelt es** dafür **Kritik***.[41]
'At 63 years old, a woman in Australia gave birth to her first child. This was greeted with harsh criticism (lit. '**it hailed criticism** for it') in Australia.'

There is no natural relation between the event of a 63-year-old woman giving birth to her first child with the meteorological event of hail coming down. Without the overt mention of the complement, the meaning of precipitation in pellets of ice prevails. Obviously, the ME of hailing comes with perils. It can be painful at the very least and destructive or even deadly in the very worst case. Thus, unsurprisingly, the metaphorical extension in (34) carries negative semantic prosody, i.e., the verb *hageln* in figurative expressions predominantly collocates with negatively evaluated words, such as *Kritik* 'criticism'. This negative semantic prosody occurs habitually with figurative expressions of hail; in fact, it seems to be the foremost sense (cf. Chapter 3, Section 3.4.3 and Chapter 4, Section 4.4). In conclusion, semasiology is also concerned with metaphorical extension, polysemy, and lexical relations.

Last but not least, lexicon-grammar interaction obviously focuses on the interaction and correlation between lexicon and grammar. As pointed out earlier in this section, the lexicon is structured by semantic and formal criteria. Likewise, constructions are form-meaning pairings (not the result of abstract syntactic rules) and thus are not fundamentally different from items in the lexicon, no matter how long or complex these word "strings" can be. For MEs, this perspective could lead to, first of all, the need to identify which word classes the German language employs for lexicalizing meteorological concepts (e.g., *Regen* [N] vs. *regnen* [V] vs. *regnerisch* [ADJ]; secondly, the need to address the syntactic constructions that are used for talking about MEs (e.g., *es gibt* + N_{ME} vs. *es* + V_{ME}); and, finally, but equally importantly, the analysis needs to show how the semantic load is partitioned between lexicon and grammar. As discussed earlier, except for in the predicate-argument type of German ME where the semantic load is somewhat shared between argument and predicate (i.e., heavy argument + heavy predicate), the predicate type presumably corresponds to a [heavy predicate + light argument]

[40] www.bild.de/regional/bremen/schwerer-orkan-fegt-ueber-bremen-8488948.bild.html (accessed April 14, 2017).
[41] www.rp-online.de/leben/gesundheit/schwangerschaft/63-jaehrige-australierin-hat-ihr-erstes-baby-bekommen-es-hagelt-kritik-aid-1.6161829 (accessed April 14, 2017).

construction, while the argument type involves to a [heavy argument + light predicate] construction. However, there are also the linguistically "weird" constructions to consider, such as the [*es gibt* + PRECIPITATE] or [*es hat* + PRECIPITATE] constructions, which are neither intransitive, strictly speaking, nor full-blown transitives either. I approach German precipitation events from the constructional perspective in Chapters 3 and 4. With regard to lexicon-grammar interaction, both onomasiology (from meaning to expression) and semasiology (from expression to meaning) play a role. The main focus of the study at hand lies in the interaction between lexicon and grammar or, more specifically, on the interaction between German ME terms and the intransitive and light transitive constructions that give them expression. In sum, focusing on grammar as the only source for typologies is overly reductionist and, ultimately, leads to superficial or, in the worst case, even erroneous classifications. The linguistic structure inherent in the lexicon must be integrated into typological approaches as well.

1.7 The structure of this dissertation

Where this introductory chapter has largely dealt with existing literature on weather events, terminology, fundamental assumptions of a variety of approaches and methods that support the present study, the remainder of this dissertation chiefly attends to the analysis of authentic data collected from the *deTenTen13* corpus and the *Das Wörterbuch der deutschen Sprache,* or *DWDS,* dictionary and corpora.

Chapter 2 details the procedures and criteria for the selection of ME constructions. I describe the corpora chosen for the data collection—mainly the *deTenTen13* corpus hosted on *SE* with additional examples from the *DWDS*, as well as *Google*; the data frame that was created for each construction; and the subsequent markup of the data frame for multiple morphosyntactic and semantic variables.

Chapters 3 contains the bulk of the quantitative and qualitative data analyses that was carried out based on an examination of 24 data frames (a data structure that organizes data into a 2-dimensional table of rows and columns), i.e., three major, unmarked constructions each with two word orders and four different precipitates. The constructions are the two argument constructions: 1. **$N_{PRECIPITATE}$ + V** (*Regen fällt* 'rain falls', *Schnee fällt* 'snow falls', *Schneeregen fällt* 'sleet falls', *Hagel fällt* 'hail falls') and 2. ***es gibt*** 'it gives'**+ $N_{PRECIPITATE}$**. The analysis involves the distribution of several syntactic and semantic variables and their absolute and relative

frequencies in order to discern patterns of constructional behaviour in interaction with the precipitate lemma that fills the argument or predicate slot of the construction. Lastly, the analyses help to explore the motivation of using different constructions for the same precipitation event.

Chapter 4 concerns a number of minor, marked constructions which German speakers also employ to talk about the weather. These constructions involve either culture-specific agentives which cause the ME, such as *Frau Holle* 'Mother Hulda' and *Petrus* 'Saint Peter' in Section 4.3.1 or prepositions, which either anchor the ME in an ongoing situation (e.g., the *am*-progressive) or which background the ME against another concomitant and profiled event (e.g., preposition *bei* 'at') in Section 4.3.2. The last section of this chapter, Section 4.3.3, again involves culturally specific items, i.e., weather compounds with *Wetter* 'weather' as the head, where the preceding modifier(s) define the type of weather, such as *Hundewetter* lit. 'dog weather' (bad weather) or *Bilderbuchwetter* 'picture book weather' (beautiful weather).

Chapter 5 provides the summary and conclusion of the dissertation. It revisits the major concepts of cognitive approaches and how these informed the study at hand. I recapitulate Eriksen et al.'s (2010) typology and how and where my findings diverge from their claims for German by detailing some of the major findings of the contextual factors each variable associates with. Furthermore, I revisit lexical typology, identify some limitations of typologies and those of my own study, and briefly survey MEs that were neglected in the dissertation at hand, such as *Wind* 'wind', *Nebel* 'fog', and *Graupel* 'graupel', which also open up avenues for future research.

CHAPTER 2
Methodology and Data

This chapter presents general information on the field of corpus linguistics, the notion of a corpus, a description of the corpora used, and the fundamentals of corpus-based linguistic analyses. The bulk of this chapter is, however, dedicated to the research design and describes the data collection procedures in a detailed manner as well as the process of data analysis in the study.

2.1 Some general remarks on corpus linguistics

To give a very general summary, corpus linguistics involves the use of digitalized language materials, usually naturally occurring, authentic material in the analysis of language. For many researchers, corpus linguistics is merely considered a methodology in order to examine a wide array of theoretically framed linguistic questions. Corpus linguistics is still a rather heterogeneous field and has no clearly defined set of methods and procedures (Biber, Conrad and Reppen 1998; O'Keeffe and McCarthy 2010; Stefanowitsch 2020), although frequently used techniques include, for example, generating frequency word lists, collocates, and concordance lines (a concordance line is a list of examples of the word or string of words searched for and as they occur in the corpus). The appeal of corpus linguistics lies in the comparative ease and speed with which the researcher can carry out machine-aided explorations and investigations of large-scale data sets, which would not be feasible to accomplish manually. The next paragraphs give some general information on features of a corpus. Section 2.1.1 elaborates on two crucial concepts of corpus design, i.e., *representativeness*, which ensures that the analyzed sampled is representative of the language or language variety under consideration, and the related notion of *balancedness*, which is a requirement for representativeness. Section 2.1.2

The notion of a corpus refers to a collection of authentic language data, usually in the form of machine-readable texts, often annotated, prepared and used for searching by a concordancer (the generic name for corpus-querying software). While the majority of corpora only include written texts, there are also corpora that consist of audio or video data that may be

used for studying multimodal communication such as gesture. Generally, corpora are accessible electronically by means of a computer. This is true particularly for languages where large collections of digital data are available. However, corpora may differ considerably in size.

Corpora for underdocumented and/or endangered languages may only exist in paper form (if they exist at all), collected by the researcher themselves, and small enough to be analyzed manually. However, Rice (2018), using data from Rice and Thunder (2017) for illustrative purposes, demonstrates how even a fairly small corpus of less than 9,000 words that is analyzed with the help of a computer can yield insights about frequency, recurring multi-word expressions, and collocates, etc., and their applications. Fillmore (1992: 35) made two observations about using corpora in linguistics analysis: "The first is that I don't think there can be any corpora, however large, that contain information about all of the areas of English lexicon and grammar that I want to explore; all that I have seen are inadequate. The second observation is that every corpus that I've had a chance to examine, however small, has taught me facts that I couldn't imagine finding out about in any other way". Obviously, since the 90s, the advent of the Internet has allowed for access to larger corpora, with billions of words, that also cover a larger variety of language(s). However, the size of a corpus does not necessarily mirror its usefulness or prolificacy, although corpus size is assumed to contribute to representativeness (cf. McEnery and Wilson 2001: 78; Biber 2006: 251), particularly with regard to "diversity", given that a broad range of varieties is included, such as genre, register, style, and mode (spoken vs. written). Corpora for well-described languages like English may be comprised of billions of words. Mark Davies' *Corpus of Contemporary American English* (COCA), for example, is a billion-word, balanced monitor corpus easily accessible for free through a web interface on *english-corpora.org*, which hosts the world's most widely used corpora. The data are comprised of spoken language, fiction, popular magazines, newspapers, and academic texts, general Web texts and blogs, and TV/movie subtitles. The same website hosts the 1994 version of the *British National* Corpus (BNC). The latest edition of the *BNC* is the *BNC XML Edition*, released in 2007.[42] The BNC is a 100-million-word text balanced corpus containing instances of written and spoken English from a wide range of genres of the late 20th century. There are, of course, also corpora for specific purposes, e.g., learner corpora, synchronic or diachronic (historical) corpora, corpora of specific dialects or varieties, such as collections of legal texts, air traffic control conversations,

[42] The corpus can be downloaded at http://www.natcorp.ox.ac.uk/ (last accessed May 30, 2024).

or acknowledgements (authors' and publishers' acknowledgements in published books or research articles). Thus, corpora may represent different modes of communication (written vs. spoken, and even including features such as gesture in video corpora). Ideally, creators of corpora should aim to make the text collections systematic, balanced, and representative so that studies may be replicable (Wynne 2005; McEnery, Xiao, and Tono 2006; Xiao 2010).

2.1.1 Two key concepts of corpus design: balancedness and representativeness

In the following, I briefly elaborate on the latter two attributes of corpus design. The two most important attributes of corpus design are balancedness and representativeness, where the first refers to the range of text categories, which are supposed to be representative of the language variety under consideration and the second refers to the extent to which a sample includes the full range of variability in a population. While neither are precisely definable or indeed achievable notions (cf. Leech 2002 and Arppe 2008: 66f.), they ought to guide corpus design. With regard to representativeness, it is a critical attribute in order to make appropriate generalizations about a language; therefore, a corpus should be as representative as possible of the language or language variety. This means that a sample of the corpus should display the full range of variability found in the respective language. According to Leech (1991: 27), "a corpus is thought to be representative of the language variety it is supposed to represent if the findings based on its contents can be generalized to the said language variety. "However, corpora often display a marked bias towards standard varieties drawn from written media such as newspapers and many studies rely completely on newspaper data. This is where the notion of balance comes into play.

A balanced corpus is one that includes a wide range of genre categories that are considered to be representative of the language or language variety examined (cf. the eight text types of genre in *COCA* mentioned above). This makes balance a somewhat vague term as the research question, i.e., the intended use, dictates what is deemed acceptable in terms of balance. Nevertheless, one of the flaws of the majority of existing corpora is that they do not contain nearly enough spoken language despite conversational speech being the primary and most basic form of language use. The *BNC*, for example, widely accepted as a gold-standard balanced corpus, contains 90% written language and 10% spoken language. According to Sinclair (2004), assuming a "neutral option", a balanced corpus would contain 50% of written and 50%

of spoken language. But if we start from the premise that most speakers are exposed to a lot more speech than writing in their daily life, Sinclair gives an estimate of up to 90% spoken language as necessary for a truly balanced corpus, which is the inverse of the BNC. A balanced corpus represents not just all parts of a variety, but is also proportionate across modality or genre categories based on the typical speaker experience.

Last but not least, a corpus study should be replicable. Replicability is key to scientific practice in order to validate research. "A result is considered replicable if a reapplication of the methods that led to it consistently produces the same result" (McEnery and Hardie 2012: 16). This process of replicating results can be done with exactly the same corpus but a different (but comparable) dataset. Replicability—or the lack thereof—has also become a topic in corpus linguistics (cf. Hartmann 2023). The *TenTen* corpora (I used the *deTenTen13* corpus) are a corpus family available in more than 40 languages, created from the Web and prepared according to the same criteria. Thus, they can be regarded as comparable and, possibly, allow for an easier replication of results.

2.1.2 Making sense of corpus data

Of course, data do not exist in a vacuum and while a lot of corpus work is descriptive, corpus linguists also interpret data against diverse theoretical claims or check it against speaker intuitions. With respect to the use of a corpus, a corpus itself does not provide meaning or function to its user. When we run a query, we have merely retrieved information—and often too much to process manually—but the data also need to be evaluated (with regard to our predictions) and interpreted. We know that a word in a given context is a function of the word's interaction with that context, or as Firth (1957: 11) famously summarised: "you shall know a word by the company it keeps" (cf. Chapter 1 of this dissertation). Thus, it is vital to look not only at "numbers in total" (frequencies of occurrence), but also at collocations.

As an example exercise pertinent to this study, searching for the simple German lemma *regnen* 'to rain', with the pre-set *Key Word in Context* (KWIC) range of three words to the left and three words to the right, gives us the top three collocates (out of twenty, and as inflected form, not lemma) as *draußen*/*Draußen* ('outside', start of sentence with capital D), and *begann* (*beginnen* 'begin' in preterite). The total numbers are given in Table 3 below. However, if we add up the inflected forms and look at the co-occurrences by lemma, the top three collocates

are *Strömen* ‘streams’, *anfangen* ‘to start’, and *draußen*/*Draußen* ‘outside’. We now know the top three collocates and their total numbers of co-occurrence as word form and as lemma. We can instantly see that *Strömen* 'streams' and *draußen* ‘outside’ are part of the top three for either column, but for word form, *begann* ‘began’ is second from the top while with regard to lemma, *anfangen* ‘start’ has an edge by quite a margin over *begann* ‘began’. This is the data, as returned by the corpus, without any meaning or function to the user. In fact, I have already carried out some additional steps by comparing the word form with the lemma. The quantitative data at hand thus provides the basis for examining the distributional differences which are, after the raw frequencies, the starting point for deriving qualitative findings from the quantitative ones. Regarding their distribution, we could ask why inflected forms of *anfangen* ‘start’ are present three times in the top twenty collocates, but *beginnen* ‘begin’ is present only once and only in its preterite form (*began* ‘began’). Moreover, the top inflected form for *anfangen* ‘start’ is either in the present perfect or past perfect (followed by the preterite, while present tense has the lowest incidence). *Beginnen* ‘begin’ and *anfangen* ‘start’, are close synonyms and a closer inspection of their semantics could help reveal the distributional differences between them. *Beginnen* ‘begin’, according to the *DWDS*, is associated with an elevated style. This would complement the result that *beginnen* only occurs in the preterite in the top collocations as the German preterite is used predominanty for written text and, accordingly, if not a manifestation of a more elevated style, then at least of a more formal genre. The present perfect is used in spoken language and less formal discourse. Furthermore, *anfangen* ‘start’, according to the *DWDS*, does indeed occur in more collocations than *beginnen* ‘begin’ and, subsequently, also has more dictionary entries.

Table 3. *Total number of collocate co-occurrences with the verb lemma* regnen *'to rain' (within a specified range of three words to the left and three words to the right of the lemma)*

Collocates		Total # of co-occurrences (word form)	Total # of co-occurrences (lemma)
*Strömen**	'streams'	7,945	7,945
begann	'began'	3,828	3,828
draußen	'outside'	3,250	4,413
Draußen	'outside'	1,163	
regnet	'rains/is raining'	2,834	3,775
regnete	'rained'	941	
aufgehört	'stopped'(.perf/.plu.perf)	2,813	2,813
angefangen	'started'(perf/.plu.perf)	2,781	
anfing	'started' (preterite)	1,908	6,024
anfängt	'start/starting' (present)	1,335	

Depending on what findings we are interested in (ideally we have formed a research question or hypothesis first, so that we derive quantitative data from qualitative input leading to more qualitative output), or even if we are just "fishing", we could investigate the difference, if there are any, in the modifying materials, e.g. the adjectives that seem to be of a mostly qualitative nature, such as intensity (*es regnete stark/heftig* 'it rained heavily') or we could check whether there are preferences of one verb over the other to occur with a particular coordinating or subordinating conjunction. The former, i.e., modifiers of quality, I investigated. The latter, i.e. verb preferences of specific conjunctions, was beyond the scope of this study. Of course, I am not so much interested in isolated words, but in multi-word expressions involving precipitate lexica in German, which make up a fuller unit of meaning relevant for how a speaker actually uses the language.

The preceding considerations gave just a small glimpse into the possibilities of discovering meaning in otherwise "meaningless" data output. To give another example, but taking into account multi-word expressions, Hunston (2002) illustrates how concordances for the phrase, *recipe for,* show how the actual use is overwhelmingly metaphorical rather than literal and, moreover, how those uses carry negative semantic prosody. That is, usually, something is a recipe for something bad (*a recipe for disaster*) rather than a recipe for food, a recipe for something positive, or a recipe for something neutral. Similarly, my data show how *es hagelt* 'it is hailing' does not usually pertain to meteorological events, but is chiefly used transitively—a formal pattern we could detect without knowing the language—and typically also carries negative semantic prosody, which refers to how a seemingly "neutral" word can take on a negative (or also positive) meaning depending on its collocate (cf. Louw 1993). The

precipitation verbs merely describe the type of precipitation, but in *es hagelte Strafen* 'it hailed punishments/penalties' hageln takes on a negative meaning. Conversely, in *es regnete Goldmedaillen* 'it rained gold medals' the verb *rain* takes on a positive meaning. The latter finding we cannot otherwise deduce necessarily without a speaker's knowledge of the language.

With respect to quantitative analysis, corpus data can be subjected to a variety of statistical analyses that can reveal findings otherwise not achievable by intuition or introspection. Some simple descriptive statistical data include frequency counts. Frequency counts can be the number of occurrences of linguistic items, as well as the number of co-occurrences of items (as shown for *regnen* 'to rain' above), for example in a certain grammatical construction. These data are usually displayed in concordance lines. Stefanowitsch and Gries (2003, 2005; and Gries and Stefanowitsch 2004) developed a number of collostructional measures, i.e., measures of association strength that "quantify how strongly (or weakly) a word and a construction are attracted to or repelled by each other" (Gries 2014: 288). Hence, these measures are more dedicated to assessing the association of a word within a syntactic construction, rather than focusing on word-to-word associations. Lastly, other more sophisticated multivariate/multifactorial methods include, for example, correspondence analysis and regression (Gries 2000, 2003; Arppe 2008). Gries (2015) argues in favour of using more mixed-effects modeling in corpus linguistics, a method of doing regression analyses more frequently used in psycholinguistics. This type of analysis allows the researcher to account for "speaker-specific and lexically specific idiosyncrasies (…) elegantly" (Gries 2015: 95). The next section describes the corpus employed in this study and the data collection process.

2.2 Data Collection

For this dissertation project, I examined meteorological events (MEs) with a focus on precipitation constructions in German. The data were obtained through *Sketch Engine* (*SE*), which is an online corpus manager and text analysis tool developed to provide users (such as corpus linguists, translators, lexicographers, language learners) with access to files from more than 800 ready-to-use corpora supported in more than 100 languages, and to allow users to create their own sub-corpora from the web. *SE* was initiated and designed by (the late) Adam Kilgarriff's company *Lexical Computing Limited*, founded in 2003. As of May 2024, there are 32

German corpora available on *SE*. Twenty-two of these are "main" corpora, i.e., they are only available to paying subscribers, and the remainder are "trial" corpora, i.e., they are available to both trial users as well as paying subscribers. I opted for the largest available corpus at the time of data collection and annotation, which was the *German Web* corpus from 2013, better known as *deTenTen13*. Although it is virtually impossible to achieve absolute representativeness in all aspects of language forms and uses, this corpus is representative of a range of language varieties thanks to its large size. Furthermore, SE offers a number of features, which support a swifter and more effortless analysis. Among others, features of *SE* include word sketches (which gave *SE* its name), word sketch differences, distributional thesauri, concordance and collocation queries, word-lists, and N-grams. A full list of its features can be found in Table 4. The most relevant and most used features used in the data analysis of this dissertation are highlighted in grey.

Table 4. *List of search and return features available in Sketch Engine*

Feature	What it does
Word Sketch	gives a one-page automatically derived summary of a word's grammatical and collocational behaviour
Word Sketch Difference	compares and contrasts two words by analysing their collocation
Distributional Thesaurus	provides automated thesaurus finding words with similar meaning or appearing in the same/similar context
Concordance Search	finds examples of a word form, lemma, phrase, tag or complex structure
Collocation Search	returns word co-occurrence analysis displaying the most frequent words (to a search word) which can be regarded as collocation candidates
Word lists	generates frequency lists which can be filtered with complex criteria
N-grams	generates frequency lists of multi-word expressions
Terminology / Keyword extraction (both monolingual and bilingual)	automatically extracts key words and multi-word terms from texts (based on frequency count and linguistic criteria)
Diachronic analysis (Trends)	detects words which undergo changes in the frequency of use in time (show trending words)
Corpus building and management	creates corpora from the Web or uploaded texts including part-of-speech tagging and lemmatization which can be used as data mining software
Parallel corpus (bilingual) facilities	looks up translation examples (EUR-Lex corpus, Europarl corpus, OPUS corpus, etc.) or builds parallel corpus from own aligned texts

The "main" corpus chosen for this study, i.e., *The corpus of the German Web 2013* or the *deTenTen13*, is described in the next section. The reasons for choosing *SE* as the corpus tool and the *deTenTen13* specifically, out of the available corpora, are elaborated on below Table 6.

2.2.1 The corpus of the German Web 2013 (deTenTen13)

In an ideal world, the source corpus for this study would be perfectly representative and balanced, reflecting the varieties of the German speech community both qualitatively and quantitatively. As this is an ideal that is impossible to achieve, I opted for the next best alternative: a corpus large enough to return sufficient examples of the phenomena of interest

that I presume to be present in the German language; that is to say, a corpus large enough to contribute to representativeness and particularly variety.

The *German WebCorpus* (hereafter *deTenTen*) is part of the *TenTen corpus family* and consists of texts gathered from the internet. The *TenTen corpus family* is a collection of web corpora in more than 40 languages, which can be accessed via *SE* (cf. Jakubíček, Kilgarrif, Kovář, Rychlý, and Suchomel 2013). Each individual language corpus has been created based on the same criteria and using the same method in order to arrive at a collection of corpora that are potentially comparable and, more importantly, comprised of "linguistically valuable web content."[43] There are four German *deTenTen corpora* available (2010, 2013, 2018, and 2020). Out of the four, I used the German *WebCorpus* 2013—or *deTenTen13*—for this study, since at the time of data collection it was the most relevant and complete. The latest two *deTenTen* corpora only became available after the data analysed here had been downloaded and annotated. The *deTenTen13* data were obtained and compiled by the software *SpiderLing* in 2013.[44] The corpus was cleaned with the *jusText* corpus tool, encoded in UTF-8, and saved as text files.[45] *SE* uses a deduplication procedure to filter and discard perfect duplicates of text as well as adapted texts. Furthermore, undesired content and spam are eliminated using specialized cleaning tools. The *deTenTen* corpus building procedure for a "linguistically valuable corpus" is described in more detail on the *SE* website (see fn.43). The creators point out that 40% to 60% of crawled data are not suitable for serious linguistic analysis due to duplicated content, unwanted content (e.g., incomplete sentences, advertisements, snippets of text), "spam" (machine-generated texts and translations as well as actual, highly disproportionate spam) and texts that are too short. These data are discarded. After cleaning, the corpora seek to have a minimum size of 10 billion words, which provided the corpus family with the name *TenTen*, i.e., ten to the power of ten or 10^{10}. The German *TenTen* corpus (two-letter ISO-639-1 language code "de") *deTenTen13* consists of 16.5 billion words of texts, which were obtained from the web in 2013 ("*deTenTen13*").[46] The corpus has been lemmatized and part-of-

[43] www.sketchengine.eu/build-a-corpus-from-the-web/ (last accessed July 25, 2019).

[44] The *SpiderLing* tool is available at https://corpus.tools/wiki/SpiderLing (last accessed May 20, 24).

[45] jusText removes undesirable content such as navigation links, advertisements, headers, footers, etc. and is available at https://corpus.tools/wiki/Justext (last accessed May 20, 24).

[46] The *German WebCorpus* 2010 (*deTen* two-letter ISO-639-1 language code Ten10) consists of 2.3 billion words. The sample corpus 2013 consists of 193,838,751 words. The corpora of the *TenTen* family are all

speech (POS) tagged using the sophisticated RFTagger *pipelinev3* developed and described by Schmid and Laws (2008). The RFTagger is particularly suitable for highly inflected languages like German and able to deal with finer-grained POS details, such as number, gender, and case, separately. A chart detailing the distribution of the parts-of-speech in the *German Web corpus 2013* can be found online[47] as well as the complete tag set.[48] The German Web corpora were crawled and processed repeatedly over the years. The composition counts of the *German Web 2013 (deTenTen13)* are given in Table 5.

Table 5. *Composition counts of the German Web 2013 (*deTenTen13*) corpus*[49]

Item	Occurences
Tokens	19,808,173,163
Words	16,526,335,416
Sentences	1,095,990,180
Paragraphs	416,650,902
Documents	49,993,180

Unfortunately, genre annotation and topic classification are not available for the *deTenTen13*.[50] Regarding the distribution of domains and subcorpora statistics (see fn.49), the majority of texts come from the German Internet *de*-domain (86.85%). Furthermore, texts were sourced from other countries where German is the official language, i.e., Austrian (*at*) and Swiss (*ch*) domains with 6.54% and 6.24% composition, respectively. Finally, there is a minimal contribution of 0.38% from "other" unspecified domains. The makeup of the subcorpora in percentages and tokens is shown in Table 6.

prefixed with their respective two-letter ISO-639-1 language code and optionally suffixed with two digits for the year of compilation.

[47] www.sketchengine.eu/detenten-german-corpus/#toggle-id-12.3.2 Corpus search (accessed July 25, 2019).

[48] www.sketchengine.eu/german-rftagger-part-of-speech-tagset/ (accessed August 10, 2019).

[49] https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fdetenten13_rft3 (last accessed August 10, 2021). As of May 2024, the composition information has been replaced with the *deTenTen20* corpus (Source: www.sketchengine.eu/detenten-german-corpus/ last accessed May 23, 2024).

[50] These data are partly available for the *German Web 2020* corpus, which also belongs to the *deTenTen* corpus family, at www.sketchengine.eu/detenten-german-corpus/ (last accessed June 4, 2024).

Table 6. *Makeup of the* deTenTen13 *subcorpora*

Subcorpora	Tokens	Percentages
Domains_.de__Germany	17,202,978,496	86.85%
Domains_of_states_with_official_language_German	19,733,653,573	99.62%
Wikipedia	4,757,965	0.02%
archaeologieforum.at__archaeology_forum	1,849,816	0.01%
links-lang.de__News_Facts_and_Action_from_McPomm	4,049,703	0.02%
mein-kummerkasten.de__Help_for_your_problems	4,851,581	0.02%
myrevelations.de__Metal_and_Rock_Online_Magazine	4,494,525	0.02%

The reasons for choosing *SE* as the corpus tool and the *deTenTen13* as the specific corpus were manifold. As stated in the introduction to this section, I wanted to ensure that I interrogated a corpus large enough to return results not just for possibly the whole range of lexical items that are part of meteorological expressions, but also all of the constructions that I assumed speakers use when talking about the weather. Second, I wanted to obtain sufficient results in terms of frequencies in order to perform at least basic statistical analyses to permit some objective estimate. Having said that, much of the analysis relies on simple inspection of data, without statistical treatment, except for some significance testing (Pearson's chi-square test) in the summaries for each precipitate type. The *deTenTen13* is (or *was*, at the point of data collection) easily the largest German corpus available, so sufficient output for both quantitative and qualitative analysis can be warranted. Third, apart from obtaining sufficient lexical items and constructions in terms of frequency, I deemed it necessary and important to find less formal and more unplanned discourse that at least resembles spoken language more closely. Many corpora are comprised of written language, specifically newspapers, fiction, scientific texts, and functional literature (e.g., the *DWDS*).[51] As pointed out above regarding "balance"

[51] The *DWDS*, too, has a designated "special" spoken language corpus ("special" because it is not balanced) with approximately 2.25 million tokens. However, this corpus is comprised of speeches (various), broadcast addresses (80hrs), excerpts from Austrian parliament protocols, excerpts from approximately 250 *Spiegel* (news magazine) interviews, excerpts from *Das Literarisches Quartett* (a TV show in which literary critics discuss new book releases), excerpts from project *Emigrantendeutsch in Israel*, and excerpts from *German Bundestag* (federal parliament) protocols. None of these genres appears promising for the purpose of extracting "weather speak" unless it is figurative use, which in itself, of course, warrants closer examination and is part of the analysis undertaken in this study. As such, the figurative use of weather expressions in more formal registers is readily apparent in many of the source files making up the *deTenTen13* corpus and will be discussed in more detail in Chapter 4.

representation, Sinclair (2005) estimates that while a neutral corpus would contain 50% of written and 50% of spoken language, for a truly balanced corpus, up to 90% spoken language is necessary.

The available German corpora offer very little with regard to "*speaking* about the weather". Nevertheless, the *deTenTen13*, as a web corpus, includes a lot of data that can be regarded as "conceptually oral" (Koch and Oesterreicher 1985, 1994, 2012). This means that the informality with which much of online discourse is realized tends to resemble spontaneous speech more than written language. Discourse from chats or message boards, for example, is much more unplanned, as opposed to political speeches which are oral but conceptually literal given that they are usually written up and highly edited before they are delivered. Much of the *deTenTen13* data thus functions as a proxy for spoken language or at least an "intermediary". In following Koch and Oesterreicher (2012/1985), I assume a continuum from "extreme orality" (e.g., an intimate conversation) to "extreme literacy" (e.g., administrative regulations). While the *deTenTen13* data is written, the medium Internet has blurred some lines (e.g., the interlocutors can be in different places around the globe in different times zones, but online at the same time) and something like an "intimate conversation"—albeit with a possibly anonymous addressee—is not only possible but probable, especially in a medium like a chat forum. Admittedly, a written dialogue does not represent the same type of spontaneous, *ad hoc* exchange as a face-to-face conversation. Nevertheless, the more informal types of weather talk presumably entail elements of "the language of immediacy" like a lower degree of planning, spontaneity, expressivity, and affective speech (cf. Koch and Oesterreicher: 2012/1985) – we shall see that this is true regarding a higher incidence of stance in combination with the PRED-Cxx. However, there is a huge variety of formality within the weather talk genre and not every weather talk situation is "conceptually oral", spontaneous, or even part of a dialogue. The *deTenTen13* domains also include more formal sources of "speaking about the weather", such as weather websites and newspapers' online publications to balance out the "conceptual orality." Moreover, *SE* allows for easy access of the concordance lines and their sources. It displays all search results and permits users to download up to 10,000 lines. By contrast, the *DWDS*, which is discussed in more detail in the last two sections of this chapter, does not display all search results (due to copyright issues) and only allows the download of up to 5,000 lines. Last but not least, while the usefulness of the web as corpus has been a matter of debate (e.g., Kilgarriff and Grefenstette 2003; Leech 2007), the *SE* corpora have not only been processed and

prepared competently by professional linguists, some are also continuously updated. In addition, *SE* comes with many tools and functionalities that allow the user to process huge amounts of data in a speedy manner.

Before elaborating on the specific queries and features offered by *SE* (cf. Table 4 above), let us briefly look at potential problems with the *deTenTen13* corpus. While the ease of replicability of a static or fixed corpus, such as the *deTenTen13*, is a plus, the creators' ambitions to improve their corpora (e.g., by removing spam or by applying better NLP tools, etc.) is beneficial to most users, but disadvantageous to those who look to perfectly replicate a study. However, the creators state that they give priority to a better corpus over perfect replicability (which is not expected in STEM fields either, cf. Jakubíček et al. 2013). As a web-crawled corpus, the *deTenTen13* comes with very little metadata. A number of texts do, of course, come with the date when the text was produced (e.g., newspaper articles, blog posts, weather forecasts, press releases, etc.), but the actual production dates of many texts cannot be reliably confirmed. In terms of region, the *deTenTen13* provides the top-level domain of the website that a text comes from and the corresponding proportion (as given above) but, of course, there is no information about which regional variety the author speaks or produces. Presumably, there is overlap between southern German dialectal varieties and varieties of German spoken in Austria and Switzerland, but this remains an empirical question when it comes to particular usage phenomena.

The next sections detail the advanced corpus searches performed in order to obtain the data for this study and describe the query syntax used to carry out those searches.

2.2.2 Searching the corpus and using CQL (the Corpus Query Language)

When working with *SE,* the first step is to pick out the respective corpus. The interface then allows for a number of actions, most of them specified in Table 4 above. As I am interested first and foremost in the use of complex weather constructions and moreover complex constructions in a natural context, I used the "advanced search" feature with the query option "CQL". *CQL*, which stands for *Corpus Query Language*, is an intuitive formal language or code used to retrieve information. It allows the user to perform searches with unspecific criteria but, more importantly, it supports queries for complex grammatical or lexical patterns. CQL is used only in *SE's* advanced concordance search and allows the user to apply search criteria which cannot be

set using the standard user interface. The CQL used in *SE* is based on the original language developed at the Corpora and Lexicons group, IMS (*Institut für Maschinelle Sprachverarbeitung* 'Institute for NLP – Natural Language Processing'), at the University of Stuttgart in the early 1990s, but varies in some ways.[52]

When searching a corpus, the results are often displayed in the form of a *Key Word in Context* (KWIC) concordance. The search string, or KWIC, that matches the search criteria is highlighted and centered with context words displayed to the left and right of it. The default number of items to the left and right in *SE* is not specified and only given as "some" (usually the user can determine the number of words that are to appear to the left and right of the KWIC). It appears to be about fifteen items to either side of the KWIC. However, *SE* allows the user to change the view from "context view" to "sentence view", which displays only the sentence containing the KWIC, but no other content of the paragraph or document. The concordance lines can be sorted alphabetically by the KWIC but also by the tokens to the left or right of the KWIC. Furthermore, it is possible to sort the concordances based on text type, i.e., the metadata or headers (such as year, source, medium, or other information about the text). Beyond being sorted, concordances can also be filtered, counted, and processed further. Further processing includes the distribution of the search word/syntax across the whole corpus. Frequencies of the search word (or other words and attributes) can be computed in the concordance. For example, the most frequent word to the left before the search carried out for Figure 3 below is *wird* (inf. *werden* 'will/going to') with 33 instances followed by the conjunction *dass* 'that', of which 28 instances introduce a subordinate clause. The top of the output file shows the CQL syntax as well as the frequency per million word count. The latter allows comparison with other corpora. The concordance lines are displayed by domain and in the order in which they are found in the corpus. Every query always starts from the beginning of the corpus, i.e., in the same order. Figure 3 shows the output of an advanced concordance query using CQL in an ungapped construction, i.e., no modifiers are allowed between the three search items: the two words *es* 'it' and *Regen* 'rain', and the lemma *geben* 'give'.

[52] The documentation in *SE* describes the CQL as implemented in *manatee 2.122* which was released in April 2015. www.sketchengine.eu/documentation/corpus-querying/ (last accessed March 19, 2018).

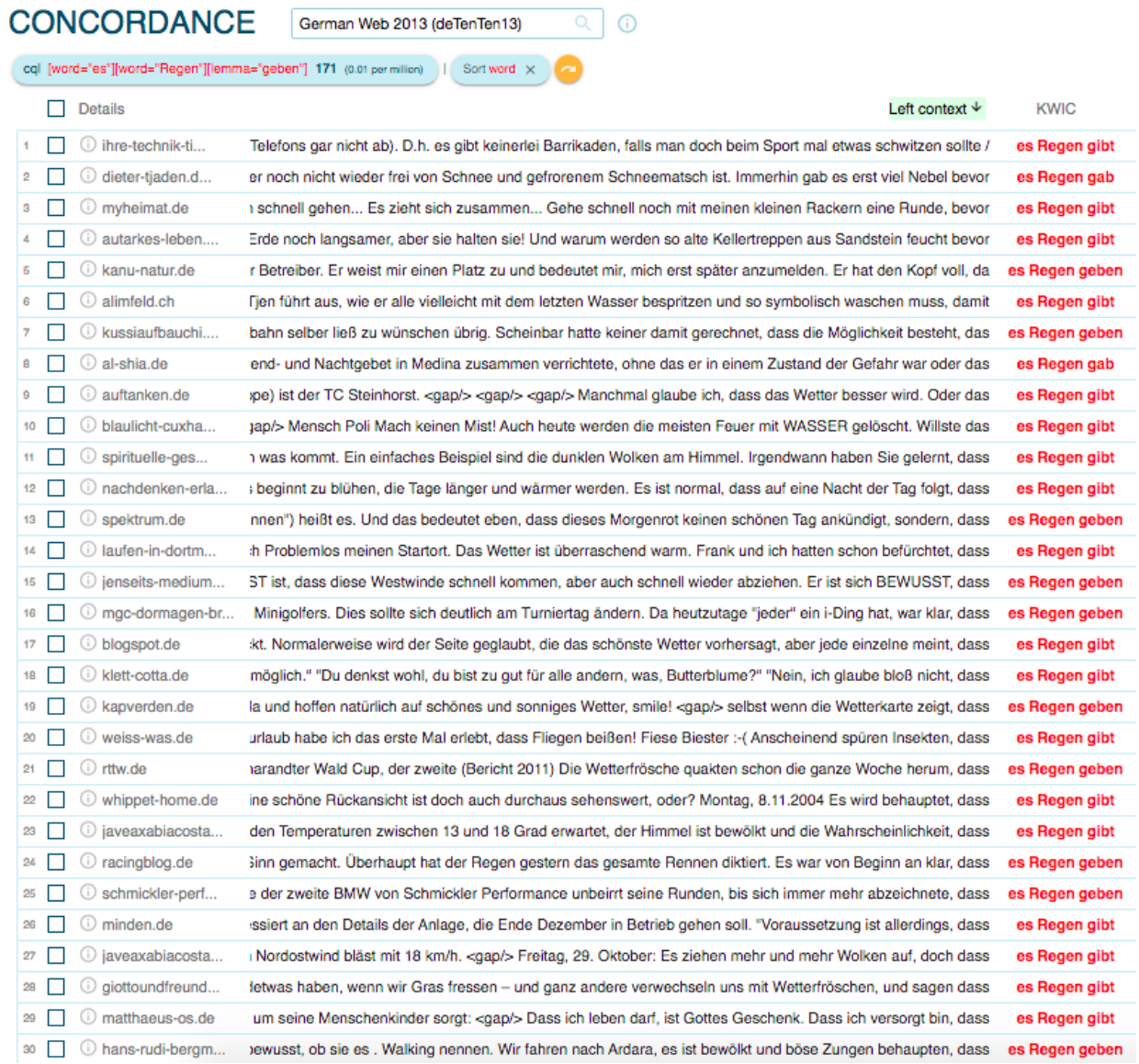

Figure 3. *Example output of CQL search for syntax [word="*es*"][word="*Regen*"][lemma="*geben*"] sorted to the left context*

CQL allows the user to set complex criteria to find one or many tokens, each of which must be inside their own pair of square brackets [ ], whereby the format is: [attribute="value"]. This means that in order to find all inflected forms of 'to rain', the format is [lemma="rain"]. In order to find the exact inflected word form 'rains', the format is [word="rains"].[53] To search for

[53] Searching for an exact word form does not distinguish between parts of speech, i.e., for English, the return "rains" could either be a verb form or the plural noun. At this point, we can make use of the POS tags that can be activated in the "view options" option or we can make the call "manually". As the German corpus is additionally tagged for case and gender, the displayed results can quickly look confusing when selecting the "view options" layout due to the plenitude of annotation information provided.

the exact phrase or string '*it is raining'*, the format is [word="it"][word="is"][word="raining"]. Empty square brackets stand for any token; that is, if we want to find instances with additional modifying material, such as '*it is always raining'* or '*it hardly rained'*, we would have to use the following format:

[word="it"] [ ][54] [lemma="rain"]

This search allows for one item (word or punctuation mark) between 'it' and 'rain', but it does not return results with more than one item between the word and the lemma. The quantifier for the number of intervening word forms has to be specified in order to delimit search returns, which is done with curly brackets. Curly brackets { } are used for repetition of the preceding token. So, if we were looking for a maximum of two lexical items in the search exemplified above, the format would be

[word="it"] [ ] {0,2} [lemma="rain"].

The length of search results can then be further delimited by using *within* and *containing* operators '< >' in order to find all tokens inside a sentence, paragraph, or document, etc. Searching for a construction with the words *geben* and *Regen*, and looking for a minimum of two and a maximum of four lexical items intervening between the verb and the noun, the query format would be:

[word="es"] [lemma="geben"] [ ] {2,4} [lemma="Regen"] within <s/>

where 's' stands for 'sentence' and the whole expression "within <s/>" stands for "being contained within a sentence" to ensure that the search result appears within the requested structure. More examples of search queries and their possible results are given in Table 7 below.

It is still difficult, if not impossible, to search corpora for constructions rather than smaller N-grams (like bigrams or trigrams)[55], as they are not (yet) annotated for constructions. Retrieving syntactic constructions such as the ones being studied in this dissertation is unworkable in an automated way. While the search needs to be delimited, I would not want to

[54] Spaces have no function in CQL so that there is no difference between [] and [ ] or [word="it"] or [word = "it"]. In this dissertation, spaces in the code are only used for easier readability. Spaces inside quotes, however, do not return results, e.g. [word=" it"], since a word cannot start with a space.

[55] In the fields of computational linguistics and probability, an N-gram is a contiguous sequence of n items from a given sample of text or speech. The items can be phonemes, syllables, letters, words or base pairs according to the application. The N-grams typically are collected from a text or speech corpus.

exclude too many returns by narrowing the gapped constructions too stringently.[56] I ran a few trial queries in order to find out what number of items between the search attributes and values would be the most appropriate and useful to search for. The trial searches confirmed that the greater the number of items that occur between the targeted words in question, the less likely they are to modify those target words. I carried out queries based on using existing annotation of words and lemmas in gapped constructions with a pre-defined number of lexical items or other modifiers (with no defined attributes, PoS tagging, etc.) between the target items. Example returns of gapped constructions by number of words in between search attributes and values are displayed in Table 7 where the syntax query was [word="es"] [lemma="geben"] [ ] {0,9} [lemma="Regen"] within <s/> .[57]

Table 7. *Examples of the* [es geben] *construction with the precipitate* Regen *'rain' allowing between 0 to 6 words in between the target items*

# of words in between the target items	Search attribute and value [word="es"] [lemma="geben"]	Examples of words in between search attribute and value	Search attribute and value [lemma="Regen"]
0	*es gab* (lit. 'it gave')	___	*Regen* 'rain'
2	*es gab*	*zumindest keinen* 'at least no'	*Regen*
3	*es gab*	*nur kurzzeitig starken* 'only briefly strong'	*Regen*
4	*es gab*	*hin und wieder leichten* 'now and again light'	*Regen*
6	*es gab*	*den für den ganzen Tag angesagten* 'the for the whole day announced'	*Regen*

By and large, the search syntax was predetermined by the syntax of German. German has verb-second (V2) word order. That is to say, in a German declarative sentence, the finite verb is always in the second position, while subject and object may be moved around. However, the predicate might be comprised of more than the finite verb and include a modal auxiliary, an infinitive, and or a past participle. If any of these elements coexist, then the inflected verb (or

[56] It is possible to use regular expressions (*RegEx*). However, during the data collection for this study, *SE* changed their interface, so that while, originally, I had relied on *RegEx* and wild cards these can no longer be used in the complex concordance search. Thus, the current data were obtained by limiting the number of items between searched attributes and their values.

[57] An example of a false return, which allows for nine items between the target words, is: *es gab weder Stühle noch Schreibtische und die Wände waren vom Regen…* 'it gave neither chairs nor desks and the walls were from the rain…'.

auxiliary) remains in V2 position while the other part of the verb complex moves to the very end of the sentence (e.g., compare the finite *Der Hund beißt den Mann* 'The dog bites the man', with *Der Hund wird den Mann beißen* 'The dog will bite the man'). This is also true for some prefixed verbs, i.e., if the finite verb has a separable prefix, the verb stem goes in second position and the separable prefix moves to a sentence final position (e.g., *Das Dorf wurde komplett* ***eingeschneit***. 'The village was completely **snowed in.'** *Durch den Schneesturm* ***schneiten*** *viele Autos auf der Autobahn* ***ein***. 'Due to the snowstorm, many cars on the highway were **snowed in**.' In subordinate clauses, the verb is also moved to the end of the sentence. This means that I had to run separate queries for V2 (*es regnete jeden Tag* 'it rained every day') and verb-final sentence structures (…, *dass es jeden Tag regnete* '…, that it rained every day'). However, due to the somewhat flexible word order in German, the verb also remains in second position when a local or temporal adverbial is in first position, as in *am Samstag schneite es* 'on Saturday it snowed', where the subject is in sentence-final position vs. *es schneite am Samstag* 'it snowed on Saturday' where the subject remains in sentence-initial position. Hence, rather than maintaining the V2/V_{final} distinction, I opted to distinguish the "word orders" based on whether the precipitate occurs before the verb (PREC1) or after the verb (PREC2) regarding the argument constructions or whether the precipitate verb occurs before the argument/subject (*regnet es* 'rains it', PREC1) or after the argument/subject (*es regnet* 'it rains', PREC2).

A fair number of lexical items can go between the subject and the verb in a subordinate clause, as in example (35), where there are seven items between the subject (*es* 'it') and the target verb (*regnete* 'rained'), although *wie aus Eimern regnen* 'come down in buckets' is, of course, highly idiomatic.

(35) *Er sagte, dass* ***es*** *noch an demselben Nachmittag* ***wie aus Eimern regnete****.*
'He said that **it rained buckets** that same afternoon.'

Nevertheless, while I restricted the number of lexical items for the queried constructions, I made a point of foregoing the tyranny of punctuation and included relevant modifying material from clauses and sentences to the left and right of the KWIC. This means that if the preceding sentence mentioned a timeframe or place that clearly modified the meteorological event in the following sentence, I included it in the annotation accordingly, since context matters. The process is further described in Section 2.2.4.

The greater number of items that appear between *es* 'it' and the target weather verb, the higher the probability that the elements searched for are not part of the same construction and that one of them forms a complex unit with some other element. The decision to limit each search to between zero and a maximum of six items between subject and verb was based on sample searches as exemplified above and also based on sample searches of the *es gibt* construction with similar results.

In (36) and (37), where a wide search was specified for ten items minimum and twenty items maximum between *es gibt* 'it gives' and *Regen* 'rain', there are nine and seven items, respectively, between *es gibt* and the target precipiate *Regen* 'rain'. The items before *Regen* 'rain' are often a list of other MEs, such as frost and hail in (36) and wind, clouds, and sun in (37), where *Regen* 'rain' is negated with *keinen* ('no_{ACC}').

(36) *Die Wachstumsperiode beträgt 8 Monate,* ***es gibt*** *im Gegensatz zu Europa keinen Frost, keinen Hagel und* ***keinen Regen*** *im Herbst.* (ingrids-welt.de 'Ingrid's world')
'The growing period is 8 months, as opposed to Europe **there is** **no** frost, no hail and no **rain** in autumn.'

(37) *Das Wetter war uns wohlgesonnen,* ***es gab*** *Wind, Wolken, zwischendurch auch etwas Sonne, aber* ***keinen Regen****.* (esv-k.de 'canoe sport')
'The weather meant well and **there was** wind, clouds, a little sun here and there but **no** **rain**.'

While the bulk of this dissertation focuses on precipitation, I carried out searches for various constructions in combination with all sorts of MEs, as detailed in Table 8.

Table 8. *Schematic CQL searches carried out for precipitate type*

Construction	precipitate/participant	CQL syntax (each two word orders)
Argument types		
[ARG + *fallen*] ARG + 'fall'	*Regen* 'rain', *Schneeregen* 'sleet', *Schnee* 'snow', *Hagel* 'hail'	[word="X"] [ ] {0,6} [lemma="fallen"] within <s/> and [lemma="fallen"] [ ] {0,6} [word="X"] within <s/>
[es + *geben* +ARG] 'it + gives + ARG'	*Regen* 'rain', *Schneeregen* 'sleet', *Schnee* 'snow', *Hagel* 'hail'	[word="es"] [ ] {0,6}[lemma="geben"] [ ] {0,6} [word="X"] within <s/> and [word="es"] [ ] {0,6} [word="X"] [ ] {0,6} [lemma="geben"] within <s/>
Predicate type		
es +V_{LEMMA} 'it' +V_{LEMMA}	*regnen* 'to rain', *schneien* 'to snow', *schneeregnen* 'to sleet', *hageln* 'to hail'	[word="es"] [ ] {0,6} [lemma="X"] within <s/> and [lemma="V"] [word="X"] within <s/>
Cognate types		
X + V_{COG_LEMMA} and V_{COG_LEMMA} + X	*Regen regnen* 'rain rain' *Schnee schneien* 'snow snow' *Hagel hageln* 'hail hail' *Schneeregen schneeregnen* 'sleet sleet'	[word="X"] [ ] {0,6} [lemma="V"] within <s/> and [lemma="V"] [ ] {0,6} [word="X"] within <s/>

The results of the queries in general are presented in Chapter 3. The exact search query is given at the beginning of each new subsection that concerns a construction type and a precipitate type, followed by the total frequency returned by the *deTenTen13*, such as

[word="es"] [ ] {0,4} [lemma="Regen"] [ ] {0,8} [lemma="geben"] within <s/> 1,658

for the [*Regen* + *geben*] construction. The cognate and the split constructions are discussed in Chapter 4. The next sections further delineate the processing of the data.

2.2.3 Extracting examples and randomizing the data

SE allows for the results to be downloaded in the following formats: .txt, .csv, .xls, .xlm, and .pdf. The output is returned in four columns, which I labeled *Reference* (the source website), *Left* (context to the left), *KWIC*, and *Right* (context to the right) respectively. In the first row of each

spreadsheet file I added the CQL code, the total number of results, and the download date, for example:

CQL [word="es"] [lemma="geben"] [ ] {0,8} [lemma="Schnee"] within <s/> 759 (0.04 per million)

SE searches the chosen corpus (or sub-corpora) methodically and serially. The concordance lines are, by default, displayed in the order in which they are found in the corpus. That is, the results are always returned in the exact same order for the exact same search. The concordances can then be shuffled and/or a random sample can be displayed by clicking on the respective icon in the task bar. The *random sample* option allows one to specify the number of concordance lines that can be downloaded (up to a maximum of 10,000 rows). The downloaded lines appear in the same order as displayed in *SE* after the randomization. The same is true for the shuffle tool. I downloaded up to 300 rows per query (or however many concordances the output consisted of if there were below 300 hits) in .xls and .csv format, so that after the data were cleaned there would be a hundred concordance lines per construction minimum, which would be sufficient for a thorough comparative and qualitative investigation. The pre-set context downloads 100 characters each to the left and right of the KWIC.

SE handles the search of corpora under a billion words in a speedy manner and returns results within seconds. Larger corpora and particularly more complex searches employing CQL containing regular expressions or frequency calculations on large corpora can take several minutes to complete. The result is displayed as the source text of the corpus and clicking on the KWIC reveals a pop-up box with even more context than given in the concordance line itself.[58] This feature turned out to be quite helpful whenever it was necessary to clarify the genre or when the KWIC with limited context was not conclusive with regard to any of the other predetermined variables. The display of lemmas, tags and other attributes can be activated using the *view options* tool, but this additional information was not needed for this study. The information icon to the left of each concordance line (or clicking directly on the source URL to the left) allows the user to display metadata about the respective concordance line. These

[58] By default, the left and right context of a KWIC concordance is determined by the width of the browser window. There is a "switch" to expand the context to 100 characters right and 100 characters left irrespective of the size of the browser window. Furthermore, it is possible to click on the KWIC itself ad reveal the "full" context, i.e., several sentences to the left and right of the KWIC depending on the length of the words and/or sentences. To give just one example where the KWIC was held stable at three words (in one specific example), there were 42 words in the displayed concordance line vs. 197 words in the pop-up context.

metadata include the token number, the document number, the first domain, the url, the document word count, the paragraph heading, the web domain, and the website. The whole concordance or just the KWIC can be displayed with their tags in an interlinearized way. Each token in *SE* is assigned a tag (also called part-of-speech tag, POS tag or morphological tag), which indicates the part of speech and grammatical category. [59] Due to the size of the corpora, the tokens are tagged automatically, which may result in some inaccuracies. However, tagging is irrelevant for the investigation carried out here, since I was, effectively, supplying my own "POS" analyses.

As the concordance lines are always returned in the same order, the returns have to be randomized, which can be done in two ways. First, after the maximum of 10,000 concordance lines have been downloaded and imported into an *Excel* spreadsheet, they can be randomized by using the 'RAND()' function. Originally, this was the only possibility for randomizing the data, since *SE* did not provide its own shuffle tool. Importing and opening the data in *Microsoft Excel*, however, would take a long time and often crash the program. A second option became available later with a new surface design in *SE*. *SE* now has a "get a random sample" and a "shuffle lines" function. Either function gives the lines in the same order as they appear in the corpus. *SE* claims that representativeness is maintained in the generated random sample. A completely different sample will be produced when changing the download value (e.g., 301 lines instead of 300). However, *SE* will generate the exact same random sample of concordance lines for a different user who searches for the exact same string and the exact same value (300 lines, or 302 lines, etc.), which is called pseudo-randomization. Using the shuffle tool on the same concordance lines more than once always produces the same new order of lines. Two different users will arrive at the same results every time they use the same search string on the same corpus and apply the shuffle tool the same number of times. *SE* designed this feature intentionally, so that queries can be reproduced with the exact same results by repeating the original search.

The next section details the construction of the data frames and the annotation of the extracted concordances.

---

[59] Cf. www.sketchengine.eu/my_keywords/tag/ and https://www.sketchengine.eu/pos-tags/ (last accessed July 21, 2019). The German RFTagger part-of-speech tagset for the *deTenTen* corpus can be found at https://www.sketchengine.eu/german-rftagger-part-of-speech-tagset/ (last accessed July 21, 2019).

2.2.4 Constructing data frames for the extracted MEs and the coding schema

The analysis in this study of German MEs, as is not unusual for corpus research, is both quantitative and qualitative in nature. For the qualitative analysis, the variety of given constructions and their interaction with the lexical items have been observed, classified, and described. The quantitative analysis is based on coded data frames for the constructions examined. For each data frame, a maximum of 300 concordance lines were downloaded where the construction queried appears in its natural context of use. Subsequently, each concordance line was manually reviewed and coded for a number of morphosyntactic and semantic features. Such an annotated data frame can then be explored with simple statistics.

The quantitative analyses will aid in exposing unique characteristics of the lexis/syntax interaction associated with the various meanings and functions of each ME construction involved in this study. The set of linguistic variables was informed by existing research (Atkins 1987; Divjak and Gries 2006; Gries 2012a; Abdulrahim 2013, etc.) where the ID tags established by those studies were deemed appropriate for the study at hand and likely to pertain to an ME schema. Some additional variables do not form part of the quantitative analysis (e.g., recurring lexical elements such as adverbs, adverbial uses, and other lexical items). Nevertheless, they are of some interest and will be referred to occasionally in the qualitative analyses in Chapter 3 and 4.

While coding the data, I went through each concordance line manually and marked the few duplicates that appeared in spite of *SE's* cleaning process as well as false returns. A false return is one in which the target precipitate/ME either was not part of the construction or merely a byproduct of it**.** This procedure has largely been described in Section 2.2.2. As stated in the preceding section, I annotated 100 random sentences for two word orders, three major constructions, and four precipitate types resulting in 24 data frames consisting of 100 lines each, i.e. 2,400 lines in total for the "default" constructions only.[60] Additionally, each line was coded for 11 variables (or ID tags), which are discussed in the following. The contextual features that

[60] The *Schneeregen* 'sleet' precipitate only returned 88 concordance lines in total for the PRED-Cxx. Furthermore, I did not annotate the many figurative *Hagel* 'hail' precipitate returns for the PRED-Cxx. However, I downloaded and annotated data for other MEs, which I originally intended to analyze, but ultimately did not mostly due to space constraints. These MEs are such as the precipitation hypernym *Niederschlag* 'precipitation' with both predicates *fallen* 'fall' and *geben* 'give', or *Nieselregen* 'drizzle' with predicate *fallen* 'fall'. Maybe these can be used in a follow up study.

each construction was coded for are nominal in nature (e.g., tense, type of modifier, genre), with other variables being categorical (binary) where the presence or absence of the feature is given as PLUS or MINUS. A feature may have more than one level. For example, tense has the levels PAST, PRESENT, FUTURE, EXISTENTIAL/GENERIC and IRREALIS, the latter expressed by the subjunctive mood or conjunctive. Both analyses, qualitative and quantitative, should aid in achieving the following objectives of this study:

1. The analyses should help us better understand the distribution of (contextual) elements in the overall data frame, allowing us insights regarding the skewed distributional patterns of the lexis/syntax interaction on a macro level, but on a micro level also within, for instance, inflectional paradigms.
2. The analyses should disclose the distribution of contextual elements for each precipitate-by-construction (also, referred to as lexis/syntax or ME/construction) interaction as a first measure towards determining variance in usage patterns concomitant to each ME.
3. Following the two previous objectives, the analysis should lastly reveal any interactional patterns among the contextual features and possible clusters of variables that are associated with certain ME uses. This critical step aids the identification of more and less prototypical uses of each precipitate within a constructional type.

In order to construct a data frame where the contextualized construction with the respective weather verb occurrence is annotated for morphosyntactic and semantic features, a number of relevant features—or variables assumed to be relevant—were chosen along the lines of *Behavioral Profiles* (or BPs; see Gries and Divjak 2006) where those variables are labeled ID tags on different levels (following Atkins 1987, cf. Gries and Divjak 2006; Gries 2012.). Arppe, in his 2008 dissertation, in which he examines common Finnish synonymous THINK verbs, makes a case for rather calling these profiles *Contextual* or *Distributional Profiles*. Table 9 below shows a selection of variables that each construction-by-precipitate was coded for, and the levels within each variable. Take note that they are not ordered (values like weight or age would be ordered), which is sometimes implied by the term *level*. The variables that I determined in my corpus are further explained below the table. Although the BP approach is geared specifically towards the examination of verb synonymy, antonymy and/or polysemy (e.g., the investigation by Divjak and Gries 2006 into Russian verbs of *trying*; Divjak and Gries 2009's investigation of

begin and *start* phrasal verbs in English and Russian; Gries and Otani's 2010 study of size adjectives in English, etc.) it can—to a degree—be applied to constructions as well.

Table 9. *Variables and variable levels for the annotation of the downloaded ME constructions analyzed in Chapters 3 and 4*

Category of variable	variable	levels
Morphological	TAM (tense/aspect/mood)	present, past, future, existential/generic, irrealis (subjunctive)
Syntactic	NEGATIVE	yes (*kein* vs. *nicht*)/ no (0)
	co-ME (co-occuring meteorological event)	co-ME involves lists of precipitates with one verb (e.g., *es gab Sonne, Wolken, Wind und Regen*)
	CLAUSE TYPE	Main clause (HS), subordinate clause (NS) and type of subordinator (e.g., *NS_ob* 'whether')
	adverbial modification (modifier)	+ (plus) / – (minus)
	Type of modifier	none (0), location (loc), temporal (temp), quantifier (quant), qualifier (qual), stance
Semantic	USE	literal / figurative
	GENRE	weather forecasts/reports weather talk meteorology literary other
	PATH	none (0), or location (L), source (S), trajectory (T), goal (G), or combinations, e.g., S+G
	COMMENT	semantic prosody (negative/positive); anything extraordinary/interesting; often an elaboration of the BYPRODUCT column, especially when there are two or more constructions in one and the same concordance line
	START/STOP	"quasi-auxiliaries" like *anfangen* 'start', *beginnen* 'begin', *aufhören* 'stop', *nicht mehr* 'no more'

TAM (tense, aspect, mood): present (including progressive), past (preterite, *Perfekt*,[61] and past perfect), future, existential/generic, irrealis (expressed by the *Konjunktiv* 'conjunctive', which is otherwise also known as subjunctive mood. To form a future sentence in German, it is possible to use the present tense in combination with a time expressions that makes it clear that the

[61] As opposed to the English present perfect, which refers to an action or state that begins in the past and continues into the present, the German *Perfekt* is usually used to refer to states and actions completed in the recent past, especially in spoken language. It can therefore often be considered a direct equivalent of the simple past tense. Because of this difference, I keep the German label *Perfekt* rather than calling it 'present perfect'.

event will happen in the future (e.g., *Morgen gehe ich ins Kino* 'I go to the cinema tomorrow'). I coded these instances as pres_FUT. The periphrastic, "proper" future tense (*Futur 1*) with *werden* 'will/to become' (*es wird regnen* 'it will rain') is marked as FUTUR or FUT. Additionally, in the analyses, I subsume the future, the irrealis, and the pres_FUT results where the incidences were low since, strictly speaking, none of these events have (yet) been actualized and belong in the realm of irrealis. (Where this is done, it is explicitly mentioned).

LITERAL VS. FIGURATIVE USE: self-explanatory.

NEGATION: in order to keep the comparability and manual coding simple, the existence of negation was coded with 0 or the different type of negation (*nicht* 'not', *kein* 'no/not a/not any').

CLAUSE TYPE: *Hauptsatz* (HS) 'main clause' vs. *Nebensatz* (NS) 'subordinate clause'. The clauses were furthermore coded according to their conjunction, e.g. NS_*dass* 'that', NS_*weil* 'because', NS_*wenn* 'when', NS_*falls* 'if', etc. and the main clauses for sentence type where there was a question (HS.Q) as in *Wird es morgen regnen*? 'Is it going to rain tomorrow?'.

MODIFIER: plus or minus for the existence or absence of a modifier, followed by:

SPATIAL MODIFIER: going back to the Talmy roots of this study, the PATH was coded for separately, i.e., where I coded for one spatial modifier ("loc") there may be several PATH components, such as a SOURCE and a GOAL (G), codes as "S+G", or a SOURCE (S) and a LOCATION (L), coded as "S+L". The PATH (P) itself can, of course, also be a component (e.g., *herunter* 'down').

NON-SPATIAL MODIFIER: *temporal* (TEMP), *quantity* (QUANT), *quality* (QUAL), and STANCE.
As noted in the previous section, modifiers that were in the context of another clause or even in the preceding or following sentence to the left or right of the KWIC were included in this variable. In a list of precipitation events from a weather forecast, for example, the place or date would of course be given at the start of the list and relate to the target construction even if it appeared in a later clause or sentence. Moreover, particularly in the more informal source material, punctuation is often randomly imposed or not given at all. Where there were several instances of one type of modifier in the same utterance, I still only coded for one modifier. Consider the following utterance in (38), which contains seven different temporal modifiers; these were still only coded as containing *a* TEMP modifier.

(38) ***Nachdem am Montag (23.04.2012)*** *bei nasskalter Witterung* ***in den frühen Morgenstunden noch Schneeregen fiel****, katapultierte uns das Wetter* ***in nur 4 Tagen*** *von einer* ***spätwinterlichen*** *Witterung in den* ***Hochsommer***! (tuttlinger-wetter.de 'weather in Tuttlingen')

'**After**[Sequence] **sleet still fell** in **the early hours of the morning**[TimeOfDay] **on Monday**[Weekday] **(April 23, 2012)**[ExactDate] in wet and cold weather, the weather catapulted us from late **winter**[Season] weather into **the height of summer**[Season] **in just 4 days**[TimePeriod]!'

Modifiers of quality and quantity can appear to be very similar and notions like *leicht* 'light' vs. *stark* 'strong/severe', which I annotated as QUAL, may also imply quantity. I decided to code modifiers that more clearly imply quantity, such as *viel* 'a lot', ein *paar Tropfen* 'a few drops' and, of course, concrete numbers (*1 liter*) and negation (*kein* 'no'), as QUANT, but modifiers that are less explicit in this regard or may also be described as having *manner* as QUAL.

co-ME (co-occurring meteorological event): where there is a list of several precipitation events that go with only one and the same verb, these were coded as co-ME as well as where other MEs, not restricted to precipitation, occurred in the previous and/or following sentence (cf. ex. (60) and (61). This is not a *yes*/*no* category, instead I coded the position in the list in which the target word occurs. For example, in a concordance like *Montag gibt es Regen und Hagel* and the target word in the *es gibt* CX query was *Hagel*, I coded *Hagel* as 'Hagel.2.' If the target word was *Regen*, I coded *Regen* as 'Regen.1.' I also coded zeugma, such as *es gab nur Regen und Schweine* 'it gave only rain and pigs'. Take note that, due to the wealth of data, I did not analyse the co-ME for every precipitate type, but only where it seemed the most pertinent, such as *Hagel* 'hail', which is a natural byproduct of other MEs to begin with and *Regen* 'rain', which is the "prototype" of precipitation events.

GENRE (weather talk, weather report, meteorology, literary, and other) Originally, this variable was coded in a much more detailed manner than turned out to be feasible, especially regarding "personal" weather talk. Thus, everything that was not discernible as "professional" weather talk (meaning the genre sub-types weather reports and forecasts, meteorology or "how weather works") was subsumed under *weather talk*. This includes personal blogs, personal travel (as opposed to travel websites that give information about the general weather at the destination at different times of the year—this would qualify as *meteorological*), events (festivals, sports events)—even if they were reported about in a newspaper, as they are not strictly speaking weather reports in a "professional" sense. *Weather forecasts* and or *reports*, which I considered to be "professional" weather talk, were also fairly apparent by the language used, by the

context (places, dates, times of day given, and usually a list of weather events), and lastly by their url (e.g., *sturmwetter* 'stormweather', *wetterprognose-wettervorhersage* 'weatherprediction-weatherforecast'). *Meteorology* is the coding used for any procedural information about how weather works, e.g., how do we get different types of precipitation or precipitation at all and also "general weather conditions" in a certain place, e.g., *winters in the North of Germany are usually rainy*, *there is no snow in Libya*. The fourth type of genre is *literary* which includes literature, poetry, fiction, audio plays/dramas, song lyrics, and MEs from religious texts. These concordances were usually obvious by the form or style they take (rhyme, literary language, choice of words, use of punctuation) and often also by their url (e.g., *hoerspielland* 'audiodramaland', *kurzgeschichten-verlag* 'shortstory-publisher'). There were only a handful of instances of "other" (e.g., *Der Name bedeutet "wo es Hagel gibt"* 'the name means "where there is hail"').

START/STOP: The *es*+PRED-Cxx revealed many co-occurences with the "quasi-auxiliaries" *beginnen* 'begin', *anfangen* 'start', and *aufhören* 'stop' (e.g., *it stopped raining*, *it began to snow*), i.e., the speakers clearly profile the incipient or final phase of the precipitation event (cf. Biber, Johansson, Leech, Conrad, and Finegan 1999). The ARG-Cxx are more likely (but not exclusively so) to include items like *nicht mehr* or *keinen mehr*, both 'no more', as in *es fiel kein Regen mehr* 'it fell no more rain'. I discuss these quasi-auxiliaries as part of the TAM-variable where applicable.

COMMENTS: The comments include information on semantic prosody (negative or positive); whether there are more than two constructions present; if there is a combination of different constructions in the KWIC (e.g., *Es regnete und Hagel fiel* "it rained and hail fell") or the same construction is repeated with different precipitates; whether there is some sort of intermediary responsible for the rain (like a deity or some other causal agentive, *Frau Holle* 'Mother Hulda', or *Petrus* 'St. Peter'); and, finally, whether anything seemed interesting or out of the ordinary**.**

Originally, I had planned a separate column for the DOMAIN (e.g., .de for German, .at for Austrian, .ch for Swiss), I later removed this coding category, as the domain is maintained in the source URL (e.g., *outdoorblogger.de*). I believe that regional differences strongly influence the use of weather constructions; however, it is usually not possible to discern whether the author belongs to a specific speaker community unless obvious lexical or informational cues were also present in the context. For the "more professional" or formal web pages (such as

newspapers), we can assume the use of *Hochdeutsch* ('high German', i.e., Standard German) which does not necessarily mirror the language, dialect or regiolect the author uses as a "private speaker". The next sections concern the *DWDS* dictionary and corpus as second source for data analysed particularly in Chapter 4 of this dissertation.

2.2.5 The DWDS *dictionary and corpora*

A second corpus, or rather a collection of corpora, was used to countercheck and confirm data obtained from the *deTenTen13* corpus: the previously mentioned *DWDS*. The *Digitales Wörterbuch der deutschen Sprache* 'Digital Dictionary of the German Language', or *DWDS*, also called *Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart* 'The word information system for the German language in the past and present' is a project of the *Berlin-Brandenburgische Akademie der Wissenschaften* 'The Berlin-Brandenburg Academy of Sciences and Humanities'. It is a digital dictionary based on the *Wörterbuch der deutschen Gegenwartssprache*, or *WDG*, 'Dictionary of contemporary German', which consists of six volumes that are linked to *DWDS's* own dictionary and text resources.[62] The *DWDS* contains several components, most notably the dictionary (including an etymological dictionary and a thesaurus), various text corpora, and statistical information. Additionally, it contains a soccer glossary, a European election glossary, and "events", which informs on the developments in soccer language in general vocabulary.[63] A list of the corpora, some of which are only accessible for registered users, is given in Table 10. Collectively, the *DWDS* corpora account for 185,800,906 documents, 4,516,364,264 sentences, and 67,412,707,131 tokens.

The search for an item in the *DWDS* dictionary returns lexicographic information from the *DWDS* electronic dictionaries at the top of the page, along with grammatical information, word formation information, the phonetic transcription, as well as a recording to hear how the word is pronounced. Furthermore, there is a 'meaning' section (*Bedeutung*), which encompasses a definition and usage examples. The etymology section (*Etymologie*) relates the word's origin, the 'thesaurus' section lists the word's common synonyms, and the *Wortprofil* 'word profile'

[62] Other sources the *DWDS* dictionary uses, such as the *Duden* and Grimm's dictionaries, can be found at www.dwds.de/d/wb-dwdswb#gwds (last accessed June 4, 2024).
[63] More information on the *DWDS* dictionary can be found at www.dwds.de/d/wb-dwdswb (last accessed June 4, 2024).

shows a word cloud of the most common collocates. Additionally, to the right of these sections, there is information about the word's frequency, the word's trajectory (from the years 1600–1999 and from 1946–2024), and a map that shows the geographical distribution based on the *ZDL* regional corpus (*ZDL* is short for *Zentrum für digitale Lexikographie der deutschen Sprache* 'Center for Digital Lexicographie of the German Language'), which contains articles from the local and regional departments of German language newspapers. Moreover, there are links to the entry of the search word in the Grimm's dictionaries and the *WDG* (s.a.) as well as a list of the *DWDS* corpora and the frequencies of the search word in each corpus. I elaborate on the statistical tools, which the *DWDS* provides, e.g., the *Wortverlaufskurve* 'word trajectory' and the *Wortprofil* 'word profile', below the table.

Table 10. *Available* DWDS *corpora*[64]

Type of corpus	description	sub-corpora
Metacorpora	multiple corpora aggregated for easy research	*Gegenwartskorpora mit freiem Zugang* 'contemporary corpora with free access'; *Historische Korpora* 'historical corpora'; *Deutsches Textarchiv Kernkorpus* + *Erweiterungen* 'German text archive core corpus and extensions'; *DWDS-Zeitungskorpus* (from 1945) 'DWDS-newspaper corpus); *ZDL-Regionalkorpus* (from 1993) 'ZDL-regional corpus'; *WebXL*
Reference corpora	corpora balanced in terms of time and text type distribution	*DWDS-Kernkorpus* (1900–1999)[65] 'core corpus 20th century'; *DWDS-Kernkorpus 21* (2000–2010) 'core corpus 21st century'; *DTA-Kernkorpus* (1598–1913) 'DTA core corpus'[66]
Newspaper corpora	either retro-digitized or purely digitally created texts from large daily and weekly newspapers	*ZDL-Regionalkorpus* (ab 1993) 'ZDL regional corpus'; *DWDS-Zeitungskorpus* (ab 1945) 'DWDS newspaper'; corpus', *Berliner Zeitung* (1945–1993) and (1994–2005) 'Berlin newspaper'; *Frankfurter Allgemeine Sonntagszeitung* (2001–2023) 'Frankfurt General Sunday newspaper'; *Frankfurter Allgemeine Zeitung* (1993–2023) Frankfurt General newspaper', *neues deutschland* (1946–1990) 'new germany', *Der Tagesspiegel* (from 1996); 'The Daily Mirror', *Die ZEIT* (1946–2023) 'The Times'
Web corpora	corpora crawled from internet-based sources	*WebXL*, *Webkorpus* 'web corpus', *Webkorpus Ballsportarten* 'web corpus ball sports', *Webmonitor*, *Jurakorpus* 'law corpus', *Medizinkorpus* 'medicine corpius', *Corona-Korpus* 'Corona corpus', *Mode- und Beauty-Blogs* 'fashion and beauty blogs', *IT Blogs*, *Blogs*
Special corpora	corpora compiled for specific genres	24 corpora in total, such as 'political speeches', 'letters from soldiers', 'film subtitles', etc.[67]

Regarding the tools for the statistical analysis of the lexicon, the *DWDS* provides *Wortverlaufskurven* 'word trajectories' (such a word trajectory is given in Section 4.2.2.1, Figure 43, for the collocation *leise rieselt der Schnee* 'softly the snow falls'), the *DWDS-Wortprofil* 'word profile', and *DiaCollo*, a software tool for the discovery, comparison, and interactive visualization of the typical word combinations for a user-specified target term. All three are

[64] A complete list of the corpora and their metadata (number of documents, sentences, and tokens) can be found at www.dwds.de/r (last accessed June 4, 2024).
[65] The *DWDS* core corpus of the 20th century is a balanced corpus with roughly 120 million tokens. It contains the following genres: fiction (28,42 %), newspapers (27,36 %), scientific texts (23,15 %), and other functional, non-fiction texts (21,05 %). It is the first reference corpus of the German language of the 20th century and claimed to be equivalent to the *BNC*, which it matches in size. The *DWDS Kernkorpus 21* is not balanced yet. Geyken (2007) provides a comprehensive overview of the *DWDS Kernkorpus*.
[66] *DTA* stands for *Deutsches Text Archiv* 'German text archive'.
[67] A complete list of the 24 corpora can be found at www.dwds.de/d/k-spezial (last accessed June 4, 2024); however, due to their specificity they were not used in this dissertation.

based on the automatic evaluation of very large corpora. The *Wortprofil* (based on the text corpora) provides information on a word's co-occurrences, frequencies, and *logDice* scores (further discussed below). On the "dictionary page", the data are presented as a visual representation in a word cloud, where the size of a collocate corresponds to the frequency of its co-occurrence, as in Figure 4, where *Eis* 'ice', *Frost* 'frost, *Kälte* 'cold', and *Regen* 'rain' are the biggest and boldest, while *fallend* 'falling', *Nebel* 'fog', and *schippen* 'to shovel', for example, are fairly small.

Figure 4. *DWDS word cloud of typical collocations of the noun* Schnee *'snow'*

A click on the *DWDS-Wortprofil* for the searched word, above the word cloud, takes the user to a separate page with detailed tabularized frequency information and *logDice* scores based on the respective grammatical function. The *logDice* score (short for the logarithm of the Dice coefficient) is a statistic measure for identifying co-occurrences, such as collocations, and expresses the typicality or strength of a collocation.[68] The *logDice* is not affected by the size of the corpus, which makes it a good measure of comparison between corpora. A weak collocation may still be of high frequency and vice versa. For example, the split construction collocates *Schnee+rieseln* 'snow+flutter', discussed in Section 4.2.2 of this dissertation, only have the fifth highest incidence (N= 2,560), but *rieseln*, a "high manner" verb, has the strongest *logDice* score out of all verb collocates with a score of 10.1. In comparison, the collocating light verb *fallen* 'fall' has an almost three times higher incidence (N= 7,512), but only a *logDice* score of 7.2. According to Rychlý, "comparing two scores, plus 1 point means twice as often collocation, plus 7 points means roughly 100 times frequent collocation [sic]" (Rychlý 2008: 9). The "theoretical

[68] The statistics of the *logDice* score are available at www.sketchengine.eu/documentation/statistics-used-in-sketch-engine/#logdice (last accessed June 5, 2024).

maximum" of the *logDice* score is 14. This applies to cases "when all occurrences of X co-occur with Y and all occurrences of Y co-occur with X. Usually the value is less than 10." (ibid.). Lastly, like the *Word Sketch Difference* function in *SE*, the *Wortprofil* allows the user to compare two lemmata and note differences or commonalities between these items.

2.2.6 The DWDS *corpus query search*

The *DWDS* corpus search allows for simple and more complex queries based on the *DCC* software and query language, such as lemma-based or exact wordforms, logical operators, such as *Regen* && *Schnee* (which returns all sentences containing the words *Regen* 'rain' and *Schnee* 'snow' in either word order), regular expressions, as well as phrases, which allow a specified number of tokens between the target items, such as "geben #<6 Schnee", which resembles the CQL query [lemma="geben"] [ ] {0,6} [lemma="Schnee"].[69] The *deTenTen13* is the main corpus used for the analysis (especially regarding Chapter 3 of this dissertation). I used the *DWDS* corpora mostly for counterchecking examples and frequencies, but they were also helpful where I tried to establish whether an item or a collocation is a neologism or has been around for a while and/or whether it has been used more formally (as in the newspaper corpora) or more informally (as in the web corpora). Keep in mind that the *deTenTen13* chiefly contains contemporary data while the *DWDS* encompasses a historical corpus as well as the *Kernkorpus* with 20th century data. I kept the *DWDS* searches simple, mostly searching for collocations, which either allowed for no gaps (*Schnee rieselt* 'snow flutters') or for the co-occurrence of two items (or more) in the same sentence (*Schnee* && *rieseln* 'snow && flutters'). The majority of *DWDS* frequencies and examples are discussed in Chapter 4. Where an example is sourced from the *DWDS*, it is indicated as such in parenthesis after the example. The same is true for examples found on *Google*. Where the source is not additionally indicated, the example stems from the *deTenTen13* corpus.

A search query can be carried out either within a specific set of corpora (cf. the first column of Table 10), but it is also possible to choose a specific sub-corpus for a search query (cf. the third column of Table 10). When searching for a simple item, like *Schneeregen* 'sleet', either

[69] The *DWDS* corpus search options and the most important search language query examples can be found at www.dwds.de/d/korpussuche#cheatsheet. The DDC query language documentation can be accessed at https://kaskade.dwds.de/~moocow/software/ddc/querydoc.html (both last accessed June 4, 2024).

in a specific corpus or a set of corpora, all of the other corpora are searched concomitantly and the results are listed on the right hand side of the page under *Belege in Korpora* 'instances in corpora'. An example is given in Figure 5, where the left box shows the frequencies returns of a search for *Schneeregen* 'sleet' in the *WebXL* corpus (bolded) and the right box lists the frequencies of a search for *schneeregnen* 'to sleet' in the *WebXL* corpus (bolded).

Figure 5. *Visual representation of frequencies of the lemmata* Schneeregen *'sleet' (left) and* schneeregnen *'to sleet' (right) in some of the corpora of the* DWDS

The main page of a corpus search shows the numbered concordance lines returned (50 per page). These are only the sentence, which contains the bolded target word, but no context where the user chooses 'full display' (*voll*). The option *maximal* generally returns the preceding and the following sentence as well. The *KWIC* display returns truncated sentences with the target word (or the first target word where there are more than one) centered. Above the concordance lines, the source is given as a clickable link that takes the user to the webpage (in case of a web corpus) or a scanned, digitized document (in case of the historical corpus, for example). The concordance lines from print newspaper generally occur with the source, i.e., the name of the newspaper and the date of the publication, but not with a link. The *KWIC* display does not occur with directly accessible links, so the user has to first open the metadata "details" by clicking on the three grey horizontal lines to the right of each concordance line. An example of a 'full' display return for the search query *Schneeregen* && *fallen* 'sleet && fall' in the *WebXL* corpus is given in Figure 6. The returns are sorted descending by date, but the *DWDS* offers 11 sorting options in total, one of which allows randomization.

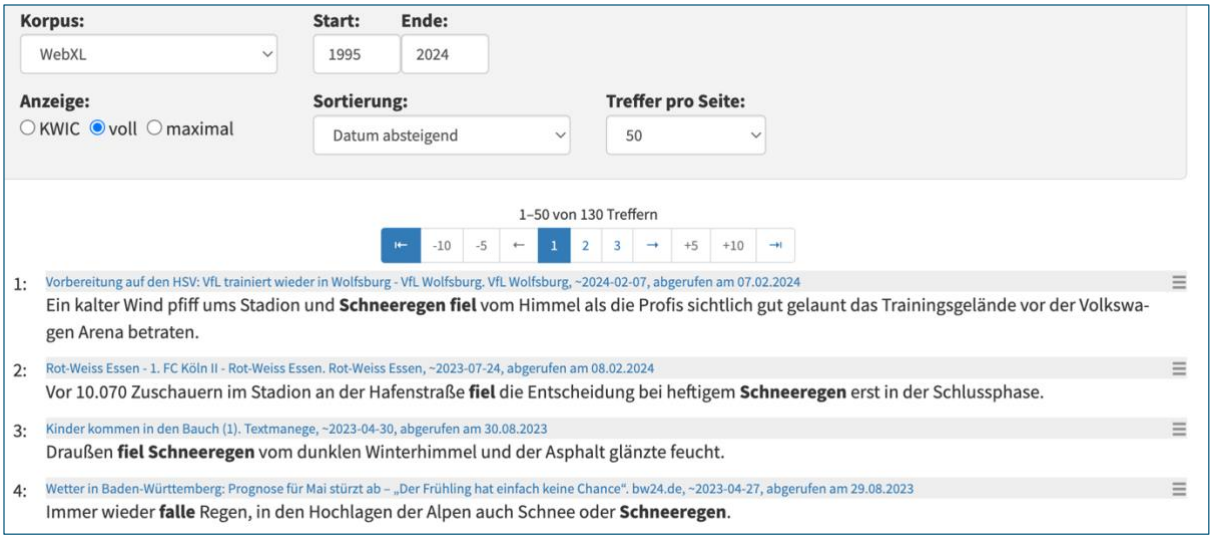

Figure 6. *Screenshot of the first four returns for the search query* Schneeregen && fallen *in the* DWDS WebXL *corpus*

As opposed to the simple word search, a search query for more than one item does not return the frequency of the returns in the other corpora, as illustrated in Figure 5. Instead, the user has to click on each corpus (listed on the right hand side) separately to see the return frequencies. Still, it is fairly easy to navigate between the different corpora and the different view options.

The search in Figure 6 also shows how false returns are part of the results, as in #2 and #4, where in the former, #2, *fallen* 'fall' is part of the phrase *Entscheidung fallen* 'make a decision', which is also the profiled event, and not *Schneeregen* 'sleet'. However, examining the ME reveals a construction, which was originally not taken into account and introduces the ME as a backgrounded event using the preposition *bei* 'at': *bei Schneeregen* 'at sleet' (cf. Section 4.3.2). In the latter concordance line, #4, *Schneeregen* is returned as the third precipitation type connected to the predicate *falle* 'fall (reportedly)' after *Regen* 'rain' and *Schnee* 'snow', where *Regen* 'rain' is the default precipitate and the latter two are restricted to the higher regions of the Alps. As with the returns from the *deTenTen13*, the results have to be manually scoured for false returns. There also appear to be quite a few duplicates between some of the corpora and within the same corpus. Due to these return overlaps and irregularities, the total frequencies given for *DWDS* searches have to be taken with a grain of salt, since I only removed those instances where low incidences were returned, which were more easily manageable by eyeballing the data. For example, the 37 returns for lemma *schneeregnen* 'to sleet' in the *WebXL* corpus (cf. Figure 5) contains one duplicate (hits #13 and #14), which only differ in their retrieval

date so that the total frequency of the predicate precipitate is 36 instances. Lastly, the returns can be exported as a CSV file or other, but I did not make use of that option.

2.3 Preview of the analyses carried out in Chapter 3 and Chapter 4

The above have described the methodology of the research as well as the corpora used for the data analysis. Chapters 3 and 4 are still very different in their approaches and analyses of the data. Chapter 3 focuses on syntactically "unmarked" constructions, i.e., the conceptually difficult participants of the ME are shoehorned into what is (mostly) a transitive sentence. The [*es*+*geben*+ME] construction is a bit of an outlier, as it basically contains two light elements: *es* 'it' and *geben* 'give'; however, it is not a full-blown transitive either. Chapter 3 involves the three "standard" constructions and their interaction with the four precipitate types, whereby the sections are organized by precipitate type. *Regen* 'rain', as the most prototypical precipitate type makes the start, followed by the *Schnee* 'snow' precipitate, then *Hagel* 'hail', and then *Schneeregen* 'sleet'. Each precipitate has a dedicated subsection, which examines the following construction types: [ARG+*fallen*], [ARG+*geben*], and [*es*+$V_{PRECIPITATE}$] (in this order).Within each construction type, I discuss each variable, as laid out in Table 9, in turn by first giving the raw and relative frequencies and then delving deeper into the semantics of each variable. For example, as demonstrated in (38), I detail the semantic and also syntactic differences within each variable, such as type of temporal or locative modification (is it deictic, does it refer to the day of the week or a season, is it a nominal or an adverbial, etc.). At the end of each precipitate section, I compare the three constructions again by addressing each variable separately. For each variable I give the raw and relative frequencies in a table. The distribution of the absolute frequencies is visualized in a vertical bar graph and the relative frequencies of each construction per variable are visualized in a horizontal bar graph. Furthermore, the summaries for each precipitate type contain the results of Pearson's chi-square (X^2) tests, which were carried out for the three-way tables, i.e., the three constructions per precipitate type, to determine whether the observed frequency distribution is due to chance. The p-value, chi-square statistics, and degrees of freedom are given for each table where the frequencies were high enough to run chi-squared tests reliably. Where the frequencies were too low for a reliable application of the chi-squared test for a significant distribution, I combined some levels of variables, such as the

peripheral paths (goal, source, path) for the spatial modifier variable. Otherwise, I report where the frequencies were too low. After all four precipitate types are discussed, I present a final summary and conclusion of all constructions and precipitates and, again, provide the relative frequencies in a horizontal bar graph. The first part of the summary and conclusion concerns the construction types and how they stack up against each other and the second part of the summary and conclusion concerns the precipitate types and how they compare. In this way, the analyses of Chapter 3 are predominantly descriptively quantitative, but the consideration of the different semantics of each variable are, of course, qualitative and provide more insight into the different usage factors.

Chapter 4 treats similarly syntactically unmarked intransitive and transitive constructions, however, the intransitives are marked in that there is no single semantically light element with regard to the first two constructions that are discussed. These constructions are the cognate construction and the split type construction where the former expresses the same ME in both the argument and the predicate (e.g., *Regen regnet* 'rain rains') while the latter expresses different facets of the ME in each component (e.g., *Schnee regnet* 'snow is raining'). As per Eriksen et al. (2010), these are the so-called argument-predicate type constructions. These constructions are also highly marked in their usage as they constitute only a minor type in German. Although they are not a default choice for everyday MEs in the German language, they are valid options with respect to specific situations and preferentially associate with certain genres. The latter is particularly true for the cognate type. I then move on to causal agentives as embodied by *Frau Holle* 'Mother Hulda' and *Petrus* 'Saint Peter', both of which are culturally specific. However, they resemble other languages, which employ causal agents as the participants responsible for the ME be it deities (weather gods, patron saints) or natural entities (clouds, the world, the sky). Furthermore, I explore construction types, which revealed themselves through the wealth of corpus data, neglected by Eriksen et al. (2010), such as the prepositional type *bei Regen* 'at rain'. This type backgrounds the ME, while another, concomitant event is profiled. Additionally, prepositions are employed in what is known as the *am*-progressive (*am* 'at the'). Lastly, I explore idiosyncratic MEs and *Wetter* 'weather' compounds, such as *Fritz-Walter-Wetter* 'Fritz-Walter-weather' or *Kaiserwetter* 'Emperor weather', which "wholesale package" whole types of weather, i.e., rainy weather and sunny weather, into singular lexical items. These compounds are, like *Frau Holle* and *Petrus*, very specific to the German culture and are rooted in historical events and figures. All of these

instances are supported by a multitude of authentic data examples drawn from the *DWDS* corpora, which, in contrast to the *deTenTen13*, allowed for quick and straightforward comparisons between different types of corpora, such as the WebXL, the newspaper corpora, and the historical corpora. In addition to the qualitative analysis of these authentic data examples. Last but not least, I also provide frequency counts from both corpora, the *deTenTen13* and the *DWDS* to corroborate my findings. While Chapter 3 is a rather dense read with a plethora of data, Chapter 4 should provide a lighter reading experience.

CHAPTER 3

Quantitative Analysis of the Target Constructions by Precipitate

The quantitative and qualitative analysis in this chapter is based on an examination of 24 [PRECIPITATE X CONSTRUCTION] data frames, i.e., three constructions each with two word orders and four different precipitates. Each data frame is comprised of 100 random concordance lines downloaded from *SE's deTenTen13* corpus, which were annotated for a variety of morphological, syntactic, and semantic variables laid out in Chapter 2. The present survey is concerned with discerning patterns of constructional behaviour in interaction with the precipitate lemma that fills the argument or predicate slot of the construction through the analysis of the distribution of said variables by considering their absolute and relative frequencies. Identifying clusters and patterns of variables and comparing their interactions will afford a more in-depth insight into the distribution and usage of weather expressions in German and aid in revealing typical sets of features that characterize the use of each construction, syntactically and semantically, in combination with a respective precipitate.

The constructions under examination are the predicate construction (PRED-Cx) *es* + V_{PREC} 'it PRECIPITATES', and the two argument constructions (ARG-Cxx): ARG_{PREC} + *fallen* 'PRECIPITATE + falls' and *es gibt* + ARG_{PREC} 'there is + PRECIPITATE', lit. 'it gives + $PRECIPITATE_{ARG}$'). The latter construction arguably contains two arguments, i.e, *es* 'it' and the precipitate itself. Accordingly, the chapter consists of four sections, one per precipitate type (*Regen* 'rain', *Schnee* 'snow', *Hagel* 'hail', and *Schneeregen* 'sleet'), with smaller subsections dedicated to each construction type per precipitate. Frequency counts and relative frequencies were determined with respect to the univariate analyses based on a number of lexico-syntactic variables. The results are then described and discussed individually for each variable and by reference to examples. Section 3.1 starts with the most frequent precipitate, *Regen* 'rain', followed by *Schnee* 'snow', *Schneeregen* 'sleet', and *Hagel* 'hail'. Each analysis will be starting with the [ARG_{PREC}+*fallen*] construction, followed by the [*es*+*geben*+ARG_{PREC}] construction and the [*es*+ V_{PREC}] construction.

3.1 Constructions with *Regen* 'rain'

3.1.1 The argument construction [Regen+fallen] *'RAIN'+'FALL'*

This section surveys the first of two ARG-Cxx in which, in contrast to the PRED-Cx, the heavy lifting of the semantic element, i.e., the respective ME precipitate, is done by the argument, while the predicate is relatively light. These two types are basically mirror images (cf. Eriksen et al. 2010) since in the PRED-Cx, the predicate encodes the ME precipitate and the argument is semantically light. Furthermore, the PRED-Cx and the ARG-Cx represent two types in a (rudimentary) typology of three constructions. The third construction, namely the ARG-PRED-Cx, is discussed in Chapter 4. Here, we are dealing with verbs *fallen* 'fall' and *geben* 'give', the first of which encodes a downward path and the second of which is often considered to be the default predicate for existential expressions (*es gibt einen Gott* 'there is a god', lit. 'it gives a god'), but is rather complex in its usage and semantics (cf. Newman 1998, 1998a) and turns out to be not necessarily the "existential default" for describing MEs in the German language. It is also notable, that the [ARG+*fallen*] constructions may employ a second light, or expletive, argument *es* 'it', which is left out, for example, where an adverbial of place or time takes its place (cf. *es fiel Regen* 'it fell rain' vs. '*am Montag/in Hamburg fiel Regen* 'on Monday/in Hamburg fell rain'). In contrast, the [ARG+*geben*] construction must employ *es* 'it' as a second argument in order to arrive at grammatically well-formed sentences. The next sections examine first the USAGE and GENRE variables for the [*Regen+fallen*] constructions and then observe the following order for the variables: TAM (tense, aspect, mood), SPATIAL, NON-SPATIAL, and the co-occurring ME (co-ME).

3.1.1.1 USAGE (LITERAL vs FIGURATIVE) and GENRE variables

This section discusses the [*Regen+fallen*] 'rain+fall' construction. The queries in the *deTenTen13* corpus were carried out using the expressions in (i) and (ii) to account for the two different word orders (PREC1 and PREC2). The (not case-sensitive) search for these two queries returned 16,860 concordance lines in total.

(i) [word="Regen"][]{0,8}[lemma="fallen"] within <s/> (N=9,682)
(ii) [lemma="fallen"][]{0,8}[word="Regen"] within <s/> (N=7,178)

Take note that the queries for this construction do not contain the second light argument *es* 'it', although *es fällt Regen* 'it falls rain', as in (39), is perfectly acceptable and used frequently. Nevertheless, this construction is only possible in the PREC2 word order configuration, as in examples (40)–(42), but not in the PREC1 order (**Regen fällt es* 'rain falls it').[70]

(39) *Am Freitag ist es an der Küste aufgelockert, ansonsten ist es stark bewölkt und **es fällt Regen***.
(donnerwetter.de 'thunderstorm' but also 'golly')
'On Friday it cleared up on the coast, otherwise it will be heavily overcast, and **it falls rain**.'

(40) *8.–11. (Februar) es wird wieder trüb, **es fällt Regen und Schnee*** (fotowetter.de 'photoweather')
'(February) 8–11 it's getting cloudy again, **it falls rain and snow**.'

(41) ***Es fällt Regen** auf den Teer. Keinem gehts jetzt besser*.
(Song *Regentage* ['rainy days'] feat. *Billy* by *Chakuza*)
'**It falls rain** onto the tar. Nobody's feeling better now.'

(42) ***Es fällt Regen**, der alles zum Wachsen bringt*. (aneirins-hallen.de 'aneirins-halls')
'**It falls rain**, which makes everything grow.'

Both word orders, PREC1 and PREC2, are possible where the light ARG is missing. Therefore, the queries were carried out for the construction not explicitly containing *es* 'it', but not explicitly excluding it from the search either. This section discusses the concordance lines returned for the queries in (i) and (ii) above. The advanced CQL query in (iii) below is merely given to show the frequency obtained from a search containing additional light ARG *es* 'it', which returned a much lower amount of concordance lines than the queries without the light ARG (N=402).[71] The search in (iv) is a simple concordance search for the exact expression with no gaps between the operators: that is to say, it does not allow for any modifying material between the three items in the search string. This query returned only 29 concordance lines in total from all of the 16.5-billion-word *deTenTen13* corpus.

(iii) [word="es"][lemma="fallen"][]{0,8}[word="Regen"] within <s/> (N=402)
(iv) *es fällt Regen* 'it falls rain' (N=29)

[70] Neither is *es fällt ein Regen* 'it falls a rain', with an indefinite article, possible without any further modification; or at least it is odd. Keep in mind that I am using PREC1 where the precipitate occurs before the argument or predicate (e.g., *Regen fällt* 'rain falls') and PREC2 where it occurs after the argument or predicate (e.g., *fällt Regen* 'falls rain').

[71] A search where the regular expression (RegEx) is not contained within a sentence, i.e. [word="es"] [lemma="fallen"] [ ] {0,8} [word="Regen"], only returns nine more concordance lines (N=411) than (iii) where the RegEx is required to be within a sentence.

Overall, 15 concordance lines were removed as false returns. Additionally, there were a total of 13 figurative concordance lines, which leaves 187 literal concordance lines for the analysis as shown in Table 11. The figurative concordance lines are mostly analogies, as in (43), and also utterances like *die Schläge fallen wie Regen* 'the blows fall like rain' and *wie ein Regen vom Himmel fällt zum dürstenden Boden* 'like rain falls from the sky to the thirsty ground'.

Table 11. *Frequencies of literal and figurative* USE *for the ARG-*fallen*-Cx and the precipitate* Regen *'rain'*

USAGE / Word order	PREC1 *Regen+fallen*	PREC2 *fallen+Regen*	Total
LITERAL	92	95	187
FIGURATIVE	8	5	13
Total	**100**	**100**	**200**

(43) *Nach dem Kahlschlag der kulturellen Revolution und der Dürre kommunistischer Gleichmacherei,* ***fiel der Regen des globalen Kapitalismus*** *mit seiner glitzernden Konsumgesellschaft hier* ***auf durstigen Boden****.* (if-blog.de)
'After the clear cut of the cultural revolution and the drought of communist egalitarianism, **the rain of global capitalism** with its glittering consumer society **fell on thirsty soil** here.'

The use of analogies means that there is a comparison to true MEs even though the ME is not actualized (which, of course, is not always the case for utterances coded as "literal" either). The distribution and frequencies of GENRE, the other main usage variable, for the [*Regen+fallen*] constructions are given in Table 12. The last column shows the relative frequency of the total per 200 concordance lines. WEATHER TALK is the principal genre type, but to a much lower degree than WEATHER TALK within the [*es+regnen*] returns. The METEOROLOGICAL genre type has the second highest incidence and makes up slightly more than one quarter of the returns. The LITERARY genre, too, is comparatively frequent at 13%. Five of the figurative instances are also LITERARY. Genre type WEATHER REPORT has the lowest incidence, a finding that will be mirrored in the [*es+regnen*] construction, but, as we will see, their frequency is proportionally much higher than that of the instances within the *es*+PRED-Cx returns. Lastly, there is one literal utterance coded as OTHER, which did not fit any of the pre-determined types of GENRE.

Table 12. *Raw and relative frequencies of the* GENRE *variable across the ARG+*fallen*-Cx and the precipitate* Regen *'rain'*

Word order/ GENRE	PREC1 *Regen+fallen*	PREC2 *fallen+Regen*	Total	Relative Frequency /200
WEATHER TALK	43	44	**87**	**43.5%**
WEATHER REPORT	7	15	**22**	**11%**
METEOROLOGY	23	28	**51**	**25.5%**
LITERARY	18	8	**26**	**13%**
OTHER	1	0	**1**	**0.5%**
FIGURATIVE	8	5	**13**	**6.5%**
Total	100	100	**200**	**100%**

The differences in distribution of the different types of genre between the two word orders are slight. WEATHER TALK and METEOROLOGY are fairly evenly distributed. However, the distributions of the LITERARY genre and WEATHER REPORTS are basically inverted in that LITERARY usages are more than twice as frequent in the PREC1 word order as in PREC2. The opposite is true for WEATHER REPORTS, which are twice as frequent in the PREC2 word than in PREC1 order. Word order differences aside, the [*Regen+fallen*] construction displays a much larger variety of genre types than the [*es+regnen*] constructions examined below (cf. Section 3.1.3) although WEATHER TALK is still the predominant genre type. Particularly noticeable is the comparatively high number of tokens for the METEOROLOGY genre, which mostly deals with explications of "how weather works" (hereafter HWW) as well as reports about the average weather (hereafter AVW) in a certain region, especially with respect to travel information. The increased proportion of METEOROLOGY tokens involves an increased incidence of EXISTENTIAL/GENERIC tokens, as detailed in the next section on the TAM variable.

3.1.1.2 The TAM (TENSE, ASPECT, MOOD) variable

The distribution and frequencies of the TAM variable for the [*Regen+fallen*] construction are given in Table 13. The last column gives the relative frequency of the total per 187 literal concordance lines, since the figurative tokens have been removed. Examining the absolute frequencies, concordance lines with an existential or generic meaning are the most frequent constituting more than one-third of the returns (N=69). EXISTENTIAL/GENERIC (hereafter EXIS_GEN) refers to a habitual use generally employed for existential accounts of MEs, such as in descriptions of how weather works, as in (46), or the general weather conditions in a specific place, as in (48) below. The existential/generic instances are even more frequent than those in the PAST tense (N=66). The category PAST tense (hereafter PAST) encompasses inflections for PAST

PROGRESSIVE (PAST_PROG), PERFEKT (PERF), and PAST PERFECT. The frequencies for the aspectual marking are not listed in the table as it did not make sense to tease them apart for the univariate analysis due to their overall low incidence. Instead, where they are present sufficiently, they are listed in a separate table and elaborated on in the running text. Where any of these TAM sub-categories are not mentioned, the default of the PAST mention is *Präteritum* 'preterite/simple past'. Regarding present tense, PROGRESSIVE (hereafter PROG) marking is listed in the table where applicable (in the same row as PRESENT and separated by a slash), as, in contrast to PAST, there is only one sub-category of aspect for PRESENT. It is striking, that out of all the PAST tokens, there are only two that also carry progressive aspect and only one that is inflected for *Perfekt* tense. The bulk of the returns are in the preterite. There are no instances of past perfect. Third most frequent, with approximately half as many tokens as the EXIS_GEN and PAST tense, is PRESENT tense (N=34), which is split up ("PRESENT / prog") between habitual present tense (PRESENT) before the slash and present progressive after the slash (prog). Without sufficient context, however, the reading is not always clear. Less than half of the PRESENT tense tokens are additionally coded for progressive aspect. Utterances involving "future intent" (i.e., concordance lines coded as FUTURE, IRREALIS, and pres_FUT) are the least frequent. It makes sense to group pres_ FUT, FUT, but also IRR tokens as "future intent" (or "irrealis"); since none of these events have yet been actualized—whether they are positively forecasted based on meteorological data or simply wished or hoped for (*könnte es bitte in Strömen regnen* 'could it please rain buckets') – – they all belong in the realm of irrealis. Having said that, teasing these three factors apart also reveals differences in usage, especially with regard to the "proper" future tense coded as FUT (meaning the periphrastic future tense *Futur 1*, as in *Regen wird fallen* 'rain is going to fall') so that the factors may be counted individually in the respective frequency tables but subsumed as "FUT_intent" in the corresponding figures, particularly in the subsections that compare all three constructions. Here, there are zero instances of overt FUTURE or IRREALIS marking. The only concordances lines that have future tense force do so via present tense inflection (coded as pres_FUT). These are concordances lines where the surface form of the verb is inflected for present tense but a temporal modifier indicates that the ME is projected to happen in the (near) future.

Table 13. *Raw and relative frequencies of the* TAM *variable across the ARG+*fallen*-Cx and the precipitate* Regen *'rain'*

Word order/ TAM level	**PREC1** *Regen fallen*	**PREC2** *fallen Regen*	**Total**	**Relative frequency** /187
PAST	26	40	**66**	**35%**
PRESENT / prog	12/9= 21	8/5= 13	**34**	**18%**
pres_FUTURE	10	8	**18**	**10%**
FUTUR	0	0	**0**	**0%**
IRREALIS	0	0	**0**	**0%**
EXIS_ GEN	35	34	**69**	**37%**
Total	92	95	**187**	**100%**
other / FIGURATIVE	8	5	13	(divided by 200) 6.5%

The returns are fairly balanced overall by worder order except for PAST tense, which is the most frequent TAM factor for PREC2 returns, but not for PREC1 returns where EXIS_GEN returns have the highest incidence. In turn, factors PRESENT and PAST are similarly frequent in the PREC1 order. There are only two concordance lines which encompass quasi-auxiliaries and both of which focus on the end-phase of the ME. What is striking, apart from EXIS_GEN having the highest incidence overall, is the lack of variety within the PAST tense tokens, especially since these predominantly occur in the WEATHER TALK genre (N=49), which assumedly exhibits a greater informality and thus "orality" than other types of genre, which would be reflected in an increased use of *Perfekt* tense.

3.1.1.3 The SPATIAL modifier variable

The distribution and frequencies of the SPATIAL modification variable for the [*Regen+fallen*] construction are given in Table 14. The last column shows the relative frequencies of the total per 187 literal concordance lines. The figurative/other uses as well as combinations are given in grey in the two rows below the "total" row. The complex SPATIAL mentions, i.e., combinations of two or more SPATIAL modifiers, were split up and counted separately. As there are six combinations of SPATIAL modification overall, the total of the relative frequencies are slightly higher than 100%. The complex SPATIAL modifications as well as the SPATIAL modifiers of the figurative utterance are listed right below the table.

***Table 14**. Raw and relative frequencies of the SPATIAL modifier variable in the ARG+*fallen *construction and the precipitate* Regen *'rain'*

Word order/ SPATIAL type	**PREC1** *Regen+fallen*	**PREC2** *fallen+Regen*	**Total**	**Relative frequency** /187
LOCATION	32	48	**80**	**43%**
GOAL	13	9	**22**	**12%**
SOURCE	8	3	**11**	**6%**
PATH	4	2	**6**	**3%**
ZERO	39	34	**73**	**39%**
Total	**96**	**96**	**192**	**103%**
FIGURATIVE	9	7	16	9%
COMBINATIONS*	4	2	6	3%

* PREC1= T+G (N=2), (L)+S (N=1), T+S (N=1); PREC2= L+T (N=1), S+(L) (N=1)

There are three results that stand out regarding the aggregate frequencies of SPATIAL modification: there is some variety of SPATIAL modification but, predominantly, the concordance lines either specify some LOCATION or no path whatsoever (ZERO), whereby LOCATIONS are slightly more frequent by four percentage points. In 30% of the returns, the LOCATION directly modifies the ME and in 13% the SETTING is provided in the larger context (the tokens were counted separately and divided by the number of literal concordance lines returned). An example of this is (44) where a couple of locations are enumerated in a separate sentence in the right context.[72] In example (45), the LOCATION is given in a separate sentence in the left context but could easily be inserted before *fiel*: *Der Regen in Bad Blankenburg* 'The rain in Bad Blankenburg'; here, there is no necessity to do so, though. The LOCATION in the left context could also function as an antecedent and the location could be referred back to as *hier* 'here': *Der Regen fiel hier so dicht*…'The rain fell so densely here…'.

(44) *Für das Wochenende erwarten Meteorologen* ***so viel Regen, wie sonst in einem ganzen Monat fällt.*** *Besonders betroffen sind die Regierungsbezirke Oberpfalz, Ober- und Mittelfranken.* (sat1bayern.de – sat1bavaria)
'For the weekend, meteorologists expect **as much rain as otherwise falls in a whole month**. The administrative districts of Upper Palatinate, Upper and Middle Franconia are particularly affected.'

[72] The wider geographical context for this weather forecast may also be guessed at from the source URL *Sat1Bayern*, knowing that *Sat1* is a free-to-air television channel and *Bayern* denotes the region of Bavaria.

(45) *Bad Blankenburg (sei) am Abend des 17. Mai "in einer Wasserwand versunken". (Thomas) Günzel: "**Der Regen fiel** **so dicht**, dass man nicht mehr über das Stadtzentrum hinwegsehen konnte."* (idea.de)
'(They say that) Bad Blankenburg was "sunk in a wall of water" on the evening of May 17th. (Thomas) Günzel: "**The rain fell** **so densely** that one could no longer see beyond the city center."

The third most frequent type of SPATIAL modification is a GOAL, which, apart from the literal utterances, occurs an additional six times occurs in a figurative utterance. Semantically, the GOAL does not essentially differ between literal and figurative sentences, as the figurative sentences mostly are analogies to actual rain events and specify how/where the rain falls (e.g., *auf die Erde*, *auf den Boden* 'onto the earth/the ground', cf. (43) above where the 'thirsty ground/soil' is metaphorical). Including figurative concordance lines, GOALS add up to 14% of the data. Other literal GOALS are both animate (experiencers) and inanimate and include: *auf unsere Böden* 'onto our soil', *ins Fahrzeug* 'into the vehicle', *zur Erde* 'to the ground/earth', *auf die Stiefmütterchen* 'onto the pansies', *aufs Vordach über ihnen* 'onto the roof above them', *ins Gesicht/auf mein Gesicht* 'into the face/on my face', *auf die MiRadler* 'onto the MiCyclists', and *auf Ihre Grünfläche* 'onto your lawn'. With respect to SOURCES, of which there are exactly half as many as GOALS, they are mostly given as *vom Himmel* 'from the sky' (N=6). Other SOURCES are, for example, *aus den Wolken*, lit. 'out of the clouds', *aus dem Stratus* 'from the stratus', *aus der geschlossenen Wolkendecke* lit. 'out of closed cloud cover'. There are three more SOURCES appearing in figurative sentences. In a similar manner to GOALS, SOURCES do not essentially differ between literal and figurative utterances and two of the figurative SOURCES are instances of *vom Himmel* 'from the sky'. Obviously, SOURCES are limited in the types they portray, as it must be "from above" as well as some synonym of *cloud*, *sky*, *heaven*, or through an animate intermediary like some weather god or similar (cf. St Peter in (80), p.132). In contrast, the types of conceivable goals are basically limitless as illustrated by the list above. Lastly, there are six instances that include a PATH. Take note that it is more likely for a PATH to occur in a combination with some other SPATIAL modifier rather than by itself. Example (46) describes the path rain takes due to the fact of gravity. Example (47) shows a combination of a PATH (*nach unten* 'downwards') with a GOAL (*auf die Erde* 'onto the earth').

(46) *(Die Schwerkraft) sorgt dafür, dass der **Regen** **abwärts und nicht aufwärts fällt**.* (importantlinks.de)
'Gravity causes **rain to fall** **downwards and not upwards**.'

(47) ***Der Regen fällt nach unten auf*** *die Erde (…).* (joerg-sieger.de)
'**The rain falls down onto** the earth.'

There are a total of seven combinations of SPATIAL modifiers, including one figurative ME, which is quite a lot. In contrast, we will see in Section 3.1.3.3 that within the [*es+regnen*] returns there are zero combinations of SPATIAL modifications. Generally, the fact that more than one-fifth of the returns contains "peripheral" SPATIAL modifiers (i.e., not a LOCATION) is striking.

Regarding LOCATIONS, other than specific places, and in line with Liégeois's (2021) findings, there is a group of LOCATIONS which refer to "cardinal" or "geographical directions", such as *im Norden* 'in the north' or *im westpazifischen Raum* 'in the western Pacific region'. Sometimes these regions are also given in a less specific manner as *in den nördlichen und zentralen Regionen* 'in the northern and central regions' or just *im Küstengebiet* 'in the coastal area' or the more general *im ganzen Land* 'throughout the country'. While the latter obviously refers to the nation where the forecast is made (in this case, Germany), there are 19 LOCATIONS (19 types, 20 tokens) which specify other continents/regions (*in Australien* 'in Australia', *in den Tropen* 'in the tropics'), countries (*in Äthiopien* 'in Ethiopia'), cities (*in Manila*), city districts (*in München-Friedenheim* 'in Munich-Friedenheim') or zoom in even further on other specific locations (*an der Fachhochschule* 'at the technical college'). Last but not least, there are the deictics, such as *hier* 'here', which sometimes combines with another SPATIAL modifier specified in the left or right context (e.g., *der Süden des Landes liegt im Winterregengebiet….. **Hier**+ME* 'the south of the country lies in the winter rain region…. **Here**+*ME'*) or *draußen* 'outside', *stellenweise* 'in places', or *verbreitet* 'widespread'. With respect to PoS, the deictics mainly consist of adverbs. While there is predominantly a nominal style for modifiers of place, deictics are merely complemented by nouns as in the bolded example above (*im Winterregengebiet….. **Hier**+ME*). The next section deals with NON-SPATIAL modifiers.

3.1.1.4 The NON-SPATIAL modifier variable (TEMPORAL, QUANTITY, QUALITY, STANCE)

The distribution and frequencies of NON-SPATIAL modifiers for the [*Regen+fallen*] construction are given in Table 15; SPATIAL mentions are listed in the last row. Only literal utterances are considered. The last column gives the relative frequency of the total per 187 concordance lines. The total amounts to more than 154% (without consideration of SPATIAL modifiers), as the majority of the data encompass more than one modifier. Overall, there are 287 NON-SPATIAL

modifiers in 187 concordance lines. Only nine concordance lines do not contain any kind of modification at all, i.e., SPATIAL or NON-SPATIAL modification.

Table 15. *Raw and relative frequencies of NON-SPATIAL modification across the ARG+*fallen *construction and the precipitate* Regen *'rain'*

Word order/ **NON-SPATIAL MODIFIERS**	**PREC1** *Regen+fallen*	**PREC2** *fallen+Regen*	**Total**	**Relative frequency** /187
TEMPORAL	52	72	**124**	**66%**
QUANTITY	37	54	**91**	**49%**
QUALITY	23	23	**46**	**25%**
STANCE	12	14	**26**	**14%**
Total	**124**	**163**	**287**	**154%**
ZERO	9	0	9	4.8%
COMBINATIONS	58	80	138	74%
SPATIAL	30	52	82	44%

Regarding the distribution and frequencies by word order (N=163 vs. N=124 respectively) for the [*Regen+fallen*] construction, the PREC2 order is more frequent by about 40 concordance lines. (Which we will see is a similar divergence as for the [*es regnen*] construction returns). The PREC2 word order does not contain a single concordance line without any modification. The most frequent modifier overall is TEMPORAL. At least one TEMPORAL modifier occurs in two-thirds of the returns. Second most frequent are modifiers of QUANTITY, which occur in roughly half of the data. This appears to be a high number for a "non-essential" or "peripheral" modifier – at least if we think of the components stipulated by the *Frame Nets*.[73] We will see in Section 3.1.3 that the [*es+regnen*] construction contains only 18 modifiers of QUANTITY in total. Modifiers of QUALITY or MANNER are present in about a quarter of all the [*Regen+fallen*] returns while expressions of STANCE have the lowest incidence in the data set at hand. This finding may go hand-in-hand with the lack of "immediacy" as expressed (or not) in the TAM variable through aspectual variation. Of course, the incidence of genre type WEATHER TALK, which may go hand-in-hand with a generally higher level of informality, is comparatively low. The following sub-sections deal with each type of NON-SPATIAL modifier one-by-one.

Modifiers of TIME. Expressions of TIME are not just the most frequent modifiers with 66% in total, they are also manifold regarding their semantics. This is true not just with respect to

[73] *Berkeley Frame Net* at framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=weather (last accessed, January 10, 2023); *German Frame Net* gsw.phil.hhu.de/framenet/frame?id=932 (last accessed, January 10, 2023)

types but also larger categories that the types can be grouped into. Again, there are more categories than suggested by Liégeois (2021) for weather reports. Furthermore, these categories seem fairly robust, i.e., they do not just contain one or two tokens, but double digits of the same type. Nine categories in total were identified (including some overlap). Two very frequent categories are "season" (*im Winter* 'in winter'; *vor allem im Frühjahr und Herbst* 'especially and in springtime and in autumn'; *ab April… kaum noch… bis es wieder Winter wird* 'hardly any… from April… until it is going to be winter again') and "months" (*im gesamten Monat Juni* 'in the whole month of June'; *bereits ab Mitte Oktober* 'already from mid of October'; *mehr als sonst üblicherweise von Mai bis Oktober* 'more than usual from May until October'). The former category is part of Liégeois's (2021) "other". Both, "seasons" and "months", appear with 13 tokens each, although some of these appear in combinations with other categories, such as "frequency". "Day of the week", which is a frequent category for weather reports, is comparatively infrequently used within the [*Regen+fallen*] data sets. There are eight tokens in total; however, there are only two, which are "stripped down" to the bare day only, i.e., *am Samstag/am Donnerstag* 'on Saturday/on Thursday'. The other six combine with other categories, such as frequency (e.g., *am Mittwoch… gelegentlich* 'on Wednesday … occasionally' or *für heute der Freitag*… *ab und zu* 'for today the Friday… every now and then'). There is, furthermore, some overlap with another category, which involves the exact date of the rainfall (N=15). Seven of these tokens mention the day of the week as well. In particular, the PREC2 word order contains 14 concordance lines specifying an actual, precise date whereas the PREC1 word order only has one instance where a specific date is identified including the year and the day of the week (*für den Sonntag 02.06.2013* 'for Sunday June 02nd, 2013') – and where, coincidentally, *no* rain is expected. As with this example, most expressions do not provide the exact date alone or the day of the week, but the time of day or frequency (e.g., *[21.7. 2011] am Vormittag …. ab und zu* '[21.7.2011] in the morning… every now and then'; *Dienstag, 28. April 2009…. besonders vormittags… zeitweise* 'Tuesday, 28 April 2009… especially in the morning…occasionally'). Generally, the exact date never occurs by itself and at least one more temporal modifier is provided as in *am 22.06.1947…. in 42 min.* 'on 22.06.1947 within 42 minutes'.

"Deictics", another frequent category in weather reports, appears with 11 tokens in the data at hand, which are all unique types (e.g., *gestern* 'yesterday', *jetzt* 'now', *damals* 'back then'). "Frequency", which Liégeois (2021) apparently consolidates with deixis, represents with

21 tokens and almost as many types (N=20). Nevertheless, regarding the latter, there is a conspicuous group of "inflections" or variations of one type (N=12), which refer to the time period of a "year". These are, for example, *jedes Jahr* 'every year', *das ganze Jahr* 'all year', *gleichmäßig aufs Jahr verteilt* 'evenly distributed throughout the year', *jährlich* 'annually', *jahrelang* 'for years', and *im Jahresdurchschnitt* 'on annual average'.

While for weather reports it makes sense that the temporal focus is chiefly on the time and day of the week, it is equally plausible that for the [*Regen+fallen*] construction overall the focus lies on larger time parcels, such as months, seasons, and, in terms of frequency, "annuality"; particularly in combination with QUANTITY (the second most frequent modifier), this construction is routinely used to report on average quantities of rainfall for a certain region, season, or duration, as is illustrated in example (48) in the next section. This finding goes hand-in-hand with the fact that more than one-fourth of the data associate with the METEOROLOGY genre. The modifiers TEMPORAL+QUANTITY appear together a total of 41 times with or without any other additional modifiers, that is roughly 30% of all concordance lines that encompass combinations. Seventeen concordance lines purely contain the combination of TEMP+QUANT. The final bigger group, "duration", only a minor category for Liégeois's (2021) weather report data, comes with 14 tokens, some of which are simple (*pro Stunde* 'per hour', *binnen 24 Stunden* 'within 24 hours') and some rather complex (*um diese Zeit von den Hundstagen bis zum Arktur* 'around this time from the dog days of summer until Arcturus'). Lastly, we can isolate the minor categories "sequence", "idiosyncratic expressions of time", and "other". While sequence appears only marginally with two tokens (*bevor* 'before' and *gleichzeitig* 'at the same time'), there are the very specific and idiosyncratic temporal modifiers, which concur with a path, like *auf seinem Weg zum See* 'on his way to the lake', *pünktlich mit der Rückkehr zum Stall* 'right on time with the return to the stable' and *auf dem White Rim Trail zur dümmsten Zeit* 'on the White Rim Trail at the most stupid time'. "Other", like the specific dates, comprises concordance lines which refer to distinct days, such as *am Metarditag* 'on Metardi day' and *St. Gallus Tag* 'St. Gallus day', which are part of rhyming weather proverbs or weather lore sayings, the so-called *Bauernregeln* lit. 'farmers' rules'.[74]

[74] *Metarditag* is on June 8th each year. There are at least five proverbs referring to the same event, one of which is *Sankt Medard keinen Regen trag', es regnet sonst wohl vierzig Tag'* (lit. 'Sankt Medard don't carry rain, otherwise it will rain for forty days'. Both spellings *Metard* and *Medard* appear to be valid). *Gallustag* is on October 16th each year and one of the proverbs referring to this day is *Wenn Sankt Gallus Regen*

Modifiers of QUANTITY. The second biggest group of modifiers is provided by expressions of QUANTITY (N=91; 49%). Scrutinizing modifiers of QUANTITY, three bigger sub-categories are apparent. As already alluded to above, there is the very factual category of quantity measured in hard numbers such as millimeters, centimeters, or liters per square meter (N=25). All of these quantifiers collocate with a TEMPORAL modifier, as they typically specify how much rain fell in what amount of time. Often, a location is provided, too, although it may be at the fringe of the left or right context and not part of the KWIC. In (48), which is a typical example of communicating the meteorological (AVW) situation for a specific place, the quantity is given in millimetres (*1800mm*). Additionally, frequency (*per Jahr* 'per year') as well as place (*Japan*) are provided.

(48) *Mit deutschem Regen sind die Niederschläge nicht zu vergleichen.* ***Insgesamt fallen in Japan 1800 mm Regen im Jahr***. (dohoo.de)
'The precipitation cannot be compared with German rain. **A total of 1800 mm of rain falls in Japan each year**.'

The other two types of QUANTITY modifiers focus either on how little rain falls or how much. It should be noted that concordance lines focusing on how little rain falls are much more productive than those which focus on how much. There are 45 concordance lines containing varieties of "little" or "no" rain, such as *wenig* '(a) little' (N=9), *ein paar Tropfen* 'a few drops' (N=4) and *etwas* 'some' (N=5), to name the most frequent ones. Less than half as many utterances (N=19) contain descriptors like *viel/sehr viel* 'a lot', *der meiste* 'the most', *große Mengen* 'large amounts' and *Unmengen* 'vast quantities'. With respect to negation ('no rain'), 16 concordance lines, almost as many as for the group containing larger amounts of rain, contain some type of negator, such as *kein* 'no' in *kein Regen* 'no rain', *gar kein Regen* 'no rain at all', *kein Tropfen Regen* 'not a drop of rain', and *meistens kein Regen* 'mostly no rain'. Obviously,

fällt, er bis Weihnachten anhält. Ist St. Gallus aber trocken, folgt ein Sommer nasser Socken. (Lit. 'If rain falls on St. Gallus, it'll last until Christmas. But if St. Gallus is dry, a summer of wet socks will follow.').

negation indicates that the quantity is zero. The "little rain" group remains the largest even if category "negation" is removed.

While TEMPORAL modifiers produced almost as many types as tokens, categories with overlap, and categories with combinations of several temporal modifiers, the QUANTITY modifiers neatly fell into three (or four) groups. To conclude, there is one instance of "other" which simply asks ***Wie viel*** *Regen fällt?* '**How much** rain falls?'.

Modifiers of QUALITY: There are 46 literal concordance lines (25%), which contain some modifiers of QUALITY or MANNER. The frequencies are evenly distributed between both word orders. The number of QUALITY modifier types is much higher than for modifiers of QUANTITY. Two "default descriptors", i.e., *leicht* 'light' und *stark* 'strong/severe', occur as the most frequent modifiers for the [*Regen+fallen*] construction. *Leichter* 'light' appears with seven tokens in total, however, *starker/stärker* 'strong/stronger' only represents with three tokens in total. A more frequent modifier here is *dicht/dichter* 'dense/more densely' (N=4), but it is further expanded upon as in *Regen fällt weiterhin als dichter Vorhang* 'rain continues to fall as a dense curtain', by coordinating a second VP: *der Regen fiel immer dichter und veranstaltete ein Trommelkonzert* 'the rain fell more and more densely and performed a drum concert', or by adding a subordinate clause, as in (45). As with the examples for *dicht/dichter* 'dense/more densely', there appears to be much more variety and creativity in relating the MANNER of the rainfall. Some examples are *moderat* 'moderatly', *fiel munter weiter* 'kept falling merrily', *45° schräg* 'at a 45° angle', *monsunartig* 'monsoonlike', *dick und schwer* 'thick and heavy'. One of the most extensive descriptions is that in (49) below, easily discernible as coming from a literary source, more specifically an award-winning narration in a newspaper.

(49) *… wenn **der Regen** anhaltender und gleichmäßiger **fiel**, bis er nurmehr ein leises Rauschen war, beinahe etwas Zärtliches, das den Schlaf und eine flüchtige und daher fast unbegreifliche Erinnerung an seine Kindheit herbeirief*, (…). (monde-diplomatique.de)
'… when **the rain fell** more persistently and more steadily, until it was just a faint rustling, almost something tender, evoking sleep and a fleeting, and therefore almost incomprehensible, memory of his childhood (…).'

The examples listed above illustrate the large variety of types found in the entirety of QUALITY modifiers: there are 34 unique types within the 54 tokens, and these do not include the variations of *dichter* 'more densly' mentioned above. The return in (50) is literally the definition of the term *Starkregen* 'heavy rain/torrential rain', where its qualities are given in terms of time and quantity: *große Mengen Regen, die in kurzer Zeit fallen* 'large quantities, which fall within a

short period of time'. The subsequent sentence defines these as "duration" and "intensity", which are, obviously, the qualities, which make *Starkregen*.

(50) *Mit **Starkregen** werden in der Meteorologie **große Mengen Regen, die in kurzer Zeit fallen**, bezeichnet. Starkregen ist somit nach seiner **Intensität** und **Dauer** definiert*. (e-hausaufgaben.de 'e-homework')
'In meteorology, **heavy rain** refers to **large amounts of rain that fall in a short time**. Heavy rain is thus defined according to its **intensity** and **duration**.'

This example is, *en passant*, an excellent reminder that features such as quality, temporality, or quantity do not have to occur as separate modifiers in German. The German language's proclivity for unhyphenated compound words is evident in its meteorological vocabulary, too. Needless to say, *Starkregen* 'heavy/torrential rain', which is part of meteorological jargon and "has a definition", does not necessarily have to be the same as *starker Regen* 'heavy/torrential rain', which any speaker may choose to use as an expression according to their own estimation of what constitutes heavy rain. However, a rain that lasts a long time is certainly a *Dauerregen* 'continuous rainfall', which is also known as *Landregen* 'steady rain', lit. 'land rain'. Again, a meteorologist will use these notions differently from non-specialists, but even within weather lexica differences in definition can be found. The *DWD*'s (*Deutscher Wetterdienst* 'German Weather Service') definition of *Dauerniederschlag* 'long-lasting precipitation', taken from their weather and climate lexicon, is given in (51). Take note that there is no entry for *Dauerregen* 'long-lasting rain', but *Landregen* 'steady rain' is defined as long-lasting precipitation in liquid form. According to *wetter.de* 'weather.de', *Landregen*, lit. 'land rain', is the designation for a long-lasting continuous rain, which seems tautological. In comparison, *wetter.de* 'weather.de' gives the definition of *Dauerregen* 'long-lasting rain' (not **Dauerniederschlag* 'long-lasting precipitation'!) as in (52), where there is no mention of quality or intensity, but merely temporality.

(51) ***Lang andauerndes Niederschlagsereignis (mehrere Stunden)** mit geringer, sich meist wenig ändernder Niederschlagsintensität und ausgedehntem Niederschlagsfeld*. (dwd.de 'German Weather Service')
'**Long-lasting precipitation event (several hours)** with low, usually little changing precipitation intensity and extensive precipitation field.'

(52) *Dauerregen:* ***Hält der Niederschlag über einen besonders langen Zeitraum an, dann spricht man von Dauerregen****. Die Zeitspanne reicht von mehreren Stunden über Tage bis hin zu Wochen****. In Europa wird in der Regel dann von Dauerregen gesprochen, wenn er mindestens sechs Stunden lang anhält****.* (wetter.de 'weather')
'Continuous rain: **If the precipitation lasts for a particularly long period of time, then it is referred to as continuous rain**. The time span ranges from several hours to days to weeks. **In Europe, continuous rain is generally spoken of when it lasts for at least six hours**.'

Either way, the fact of the matter is that these compounds, which are used widely by experts and lay people, often encode the defining features of the rain in the first compound, which makes the use of additional modifiers redundant. Other than *Starkregen*, *Dauer-* and *Landregen*, there are also *Platzregen* 'downpour', *Wolkenbruch* 'cloudburst/downpour', *Sprühregen* 'fine rain', lit. 'spray rain', and *Nieselregen* 'drizzle', to name a few. Speakers are unlikely to turn most of these compounds into verbs (**es starkregnet*, **es platzregnet*, **es landregnet*). The exception is noun *Nieselregen* 'drizzle' where the verb *nieseln* 'drizzle' is well established and predates the compound noun.[75] However, just like verb *schneeregnen* 'sleet', there are verbs, which are used and attested, but not part of the *Duden*'s inventory, such as *dauerregnen* 'to rain continuously' in (53) and *sprühregnen* lit. 'sprayrain' in (54). For the latter, the *deTenTen13* only returned 10 concordance lines in total. The search for these compound verbs was carried out separately. They were not part of the 200 annotated PRED returns.

(53) *Zudem ist der Regen regional unterschiedlich verteilt: während* ***es*** *auf Phuket* ***dauerregnet****, kann Chiang Mai tagelang sonnig sein – oder umgedreht*. (thaiminator.de)
'In addition, the rain is distributed differently regionally: while **it rains continuously** on Phuket, Chiang Mai can be sunny for days – or vice versa.'

(54) *Ganz, ganz schlechtes Wetter. Wenn es wenigstens richtig regnen würde, aber nein,* ***es sprühregnet*** *so vor sich hin und hüllt den Tag in ein freundliches Dunkelgrau*. (gudrun-plett-blog.blogspot.co.at/)
'Very, very bad weather. If only it would rain properly, but no, **it is sprayraining** away and cloaks the day in a friendly dark grey.'

In the *deTenTen13* corpus, there are practically as many returns for *dauerregnen* 'to rain continuously' (N=89) as for *schneeregnen* 'to sleet' (N=88). Be that as it may, while all of the *schneeregnen* 'to sleet' returns are literal, this does not seem to be the case for *dauerregnen*. Nonetheless, the crux of the matter is that where the aforementioned compound nouns are

[75] The verb *nieseln* 'to drizzle' and the compound *Rieselregen* 'trickle rain', which largely has fallen out of use, date back to the 18th century according to the *DWDS's* etymological dictionary. It is possible that *Nieselregen* 'drizzle rain' is a variation of the *anlaut*. This is also true for *rieseln* and *nieseln* (cf. Chapter 4, Section 4.2.2).

used, they will appear in an ARG-Cx and not in the ostensibly "default" *es*+PRED-Cx. The question, then, is by whom these compounds are used in the first place and whether the context or genre dictates their usage. The usage of these compound nouns is further discussed in Chapter 4.

Modifiers of STANCE. The NON-SPATIAL modifier type with the overall lowest frequency, STANCE, returned 26 concordance lines in total (14%), which are fairly evenly distributed between the two word orders. There are eight instances of positive semantic prosody, with half of these containing some negation (e.g., *zum Glück fiel gar kein Regen* 'luckily, no rain fell at all'; *Das Wetter meinte es gut mit uns, im Laufe des Vormittages fiel kein Tropfen Regen* 'The weather was kind to us, not a drop of rain fell during the morning'). Positive semantic prosody without negation exists where rain is necessary for plants and fruit to thrive, as in the following example:

(55) ***Fällt Regen*** *auf Ihre Grünfläche,* ***so ist dies gut für Ihre Pflanzen***. (gaertnerei-welling.de 'garden center-Welling').
'**When rain falls** on your lawn, **it is good for your plants**.'

There are two instances in which both positive and negative semantic prosody are present: the speaker does not enjoy the rain but acknowledges that it is good for the environment and/or that other people are glad about the rain. Negative semantic prosody is mostly about too much rain falling and negatively impacting events (e.g., an outdoor concert) and activities (e.g., a bike ride). In a few cases, the speaker laments the lack of rain as in an utterance where the speaker does not have an "excuse" to spend time inside a museum. Overall, where STANCE is given, it is not overly emphatic or expressed as strongly (cf. the *es*+PRED-Cxx in Section 3.1.3), but in a more subtle manner as in (56), where it is mostly expressed through the negative semantic prosody of adverb *jedoch* 'though/yet/however' as well as the verb *beeinträchtigen* 'affect'.

(56) ***Der Regen fällt jedoch nur sehr sporadisch*** *und ist meist örtlich begrenzt, so dass das Reisen nicht wesentlich beeinträchtigt wird*. (afrikaaktuell.de 'africauptodate').
'**The rain falls very sporadically though** and mostly localized, so that travelling is not significantly affected.'

The returned concordance line in (57) represents one of the strongest expressions of STANCE. While *ich ahne Böses* (lit. 'I anticipate bad things') is evidently negative, the last two sentences display some sort of indifference or resignation, to perhaps make the best of the

situation. The source URL suggests that the author is a flyfisher on a fishing trip that turned out to have less than ideal meteorological circumstances.

(57) *Um ca. 07.00h erwache **ich** und **ahne Böses**. **Es windet sehr stark und gleichzeitig <u>fällt</u> heftiger <u>Regen</u>**. Heute eilt es definitiv nicht. Wir frühstücken sehr gemütlich.....* (flyfisherman.ch)
'I wake up around 7am with **misgivings**. **It is very windy and heavy <u>rain is falling</u> at the same time**. We are definitely not in a rush today. We have a very leisurely breakfast…'

The examples in (56) and (57) also differ in GENRE. Example (56) is evidently METEOROLOGICAL (one of three concordance lines) and deals with the general weather (AVW) in some region of Africa. Example (57) is clearly more informal WEATHER TALK (one of 20 tokens), which is more prone to contain expressions of STANCE, as we have seen in the section on the [*es+regnen*] construction. In four other cases, STANCE is related in LITERARY utterances, as in (59). What the examples above have in common is that they are two of the few (six in total) that contain a QUALITY modifier (i.e., *sporadisch* 'sporadically', which also denotes frequency, and *heftig* 'severe'). It seems intuitive for modifiers of STANCE and QUALITY to go hand-in-hand, considering that many expressions of QUALITY/MANNER contain semantic prosody (e.g., *destructive/devastating rainfall*, *much needed rain*); however, this is not the case for the [*Regen+fallen*] constructions, at least not with respect to combinations of modifiers. The most frequently combined modifiers are STANCE+TEMPORAL (N=17, seven of these additionally contain a QUANTITY modifier), which seems trivial considering that TEMPORAL is the most frequent modifier overall. Another 11 times, STANCE collocates with a modifier of QUANTITY, as in (58) below, where the quantity is zero (*kein* 'no'). Another 11 times, STANCE combines with a SPATIAL modifier, as in (56) above (*örtlich begrenzt* 'localized') and in (59) below (*auf mein Gesicht* 'on my face').

(58) *Man **leidet mit**, **wenn** 10 Jahre **kein <u>Regen fällt</u>**.* (die-leselust.de 'the-readinglust')
'One **suffers along** when **no <u>rain falls</u>** for 10 years.'

(59) ***<u>Der Regen fällt</u> auf mein Gesicht** und **es fühlt sich gut an**, der Hund bellt*.
(groessenwahn.blogspot.de 'delusions-of-grandeur.blogspot')
'**<u>The rain is falling</u> on my face**, and **it feels good**, the dog is barking.'

Example (57) above does not just exemplify stancetaking but also nicely illustrates the combination of two different MEs by employing two constructions, a *es*+PRED-Cx (*es windet stark* lit. 'it is winding severely') plus the topic of discussion in this subsection, the ARG-Cx *Regen+fällt* 'rain is falling' (rather than doubling up on the *es*+PRED-Cx: *es windet und regnet stark* 'it is winding and raining severely'). This is just one type of co-ME that occurs in the [*Regen+fallen*] data sets and is discussed further in the following section.

3.1.1.5 The co-ME (CO-OCCURRING meteorological events) variable

Lastly, regarding the co-ME variable, less than half of the returned [*Regen+fallen*] concordance lines (48%), do not contain any mention of other MEs. The numbers are given in Table 16. "Multiple MEs" encompasses comma-separated lists and coordinated MEs where *Regen* 'rain' is either the first or second ME of several. There are only seven instances in total in the returns for the [*Regen+fallen*] constructions. The majority of co-MEs appear in a previous clause (N=47 or 25% of the returns), and 27 co-MEs (14%) appear in the following clause; in 15 tokens (8%) there are MEs in both the previous and the following clause or sentence.

Table 16. *Raw and relative frequencies of the co-ME variable across the ARG+*fallen*-Cx and the precipitate* Regen *'rain'*

Word order / co-ME	**PREC1** *Regen+fallen*	**PREC2** *fallen+Regen*	**Total**	**Relative Frequency** /187
ZERO	50	39	89	**48%**
Multiple MEs in same clause	6	1	7	**4%**
Previous clause	21	26	47	**25%**
Following clause	12	15	27	**14%**
Previous & following clause	3	13	16	**8%**
Other	0	1	1	**1%**
Total	**92**	**95**	**187**	**100%**

Out of the 74 concordance lines that contain an ME in either the previous or the following clause (N=47 and N=27), 33 relate to temperature, as shown in the examples below, where the temperature is given in in the preceding clause, as in (60), and in the following clause, as in (61), where the target ME furthermore is negated.

(60) *In den Tropen* ***liegt die Durchschnittstemperatur bei 27°C****. Gleichzeitig* ***fällt sehr viel Regen****.* (luftrettung-hamburg.de 'air-rescue-hamburg')
'**The average temperature** in the tropics **is around 27°C**. At the same time, **a lot of rain falls**.'

(61) *In dieser Zeit* ***fällt kein Regen*** *und* ***die Temperatur steigt***. (erlebe-thailand.de 'experience-thailand')
'During this time **no rain falls** and **the temperatures rises**.'

Unsurprisingly, the majority of these co-MEs associate within the METEOROLOGICAL genre (N=15) and WEATHER REPORTS (N=11); only seven tokens are part of WEATHER TALK. While the first two types of genre undoubtedly strive for a completeness of reporting on the current or average weather, WEATHER TALK, in its more informal manifestation, is under no such obligation. It seems

reasonable for speakers to only additionally report on temperature where it is either relevant or particularly salient. This paragraph concludes the discussion of the [*Regen+fallen*] construction and its variables. The next subsection explores the distribution and frequencies of the second ARG-construction in combination with the precipitate *Regen* 'rain', [*Regen+geben*].

3.1.2 The argument construction [Regen+geben]*, lit. 'rain + V. 'give' ('there is rain')*

This section discusses the [*Regen+geben*] construction for both PREC1 and PREC2 word orders. The queries in the *deTenTen13* were carried out using the expressions in (v) and (vi) to account for the different word orders. There are 2,915 returns in total for both queries (1,257 returns for PREC2 and 1,658 returns for PREC1). Originally, I carried out a search in a similar manner to the [ARG+*fallen*] search, i.e., without specifying the light ARG *es* 'it'. However, since *es* 'it' is obligatory in this construction, the results for either construction returned very similar numbers (cf. fn.76). I therefore opted to analyze the data containing *es* 'it'.[76]

(v) [word="es"][]{0,4}[lemma="Regen"][]{0,8}[lemma="geben"] within <s/> (N=1,658)
(vi) [word="es"][]{0,4}[lemma="geben"][]{0,8}[lemma="Regen"] within <s/> (N=1,257)

Twenty false returns (regarding word order) and one double entry had to be removed from the first 200 concordance lines and were replaced with the correct word order. All constructions are literal; there is not a single concordance that is used figuratively. In the following subsections, I report on the results of the various variables I coded in the construction data frames. The next section deals with variable GENRE.

3.1.2.1 USAGE (LITERAL VS FIGURATIVE) AND GENRE VARIABLES

The distribution and frequencies of the variable GENRE for the [ARG-*geben*] construction in combination with precipitate *Regen* 'rain' are shown in Table 17. The last column gives the relative frequency of the total per 200 concordance lines. As already noted, there is not a single

[76] The search for query "[lemma="geben"][]{0,8}[word="Regen"] within <s/>" returned 1,293 concordance lines and the search for query "[word="Regen"][]{0,8}[lemma="geben"] within <s/>" returned 1,704 concordance lines. Total number of results adds up to 2,997, which is a difference of 82 results compared to the queries in (vii) and (viii).

return where the weather event is non-literal or "other". With respect to the total counts for the variable GENRE, WEATHER TALK is the principal genre with roughly two-thirds of all returns, which—as we will see— puts this construction in the middle of the [*es*+PRED] construction and [*Regen+fallen*] returns, frequencywise. The second most frequent type of genre is LITERARY (N=30), closely followed by WEATHER REPORTS (N=25). Lastly, there are 15 tokens from the METEOROLOGICAL genre.

Table 17. *Distribution and frequencies of the GENRE variable across the ARG+*geben*-Cx and the precipitate* Regen *'rain'*

Word order/ GENRE	**PREC1** *Regen+geben*	**PREC2** *geben+Regen*	**Total**	**Relative frequency** /200
WEATHER TALK	55	75	**130**	**65%**
WEATHER REPORT	19	6	**25**	**12.5%**
METEOROLOGY	8	7	**15**	**7.5%**
LITERARY	18	12	**30**	**15%**
OTHER/FIGURATIVE	0	0	**0**	**0%**
Total	**100**	**100**	**200**	**100%**

With respect to word order, the frequencies are somewhat evenly distributed. The most pronounced difference shows up in WEATHER TALK, where the PREC2 order is one-third more frequent than PREC1. By contrast, the PREC1 order contains one-third more tokens from LITERARY sources as well as three times as many instances of WEATHER REPORTS. The predominance of WEATHER REPORTS in the PREC1 order is explained by the presence of modal verbs (N=15, mostly *kann* 'can') and/or other finite auxiliaries to form the periphrastic future tense (*werden* 'will, N=4). The LITERARY tokens come from a variety of sources: religious texts, song lyrics, fiction, and a proverb. Interestingly, almost one third of the concordance lines contain some sort of negation. With respect to TAM factors, eight are in the periphrastic future tense and five represent "true" IRREALIS cases, as in (62), which contains negation *niemals* 'never' and is a line from a song for children called *Hurra, es regnet* 'Hooray, it's raining'. Obviously, in addition to being marked as IRREALIS, the meaning is existential as it references the existence of rain.

(62) *Stell dir vor,* ***es gäbe niemals Regen****, Tag für Tag nur Sonnenschein.* (rainerherzog.de 'Rainer Herzog')
'Imagine **it would never give rain**, only sunshine day after day.'

Scrutinizing the sources of the WEATHER TALK data points further reveals that thirty-one are from private or personal blogs: these mostly deal with the topics of travel and sports. The TAM factors, which were already touched on above, are further discussed in the next section.

3.1.2.2 *The TAM (TENSE, ASPECT, MOOD) variable*

The distribution and frequencies of TAM variables for the [*Regen+geben*] construction are given in Table 18. The last column shows the relative frequency of the total per 200 concordance lines (no instances were removed for being figurative). Examining the total counts, concordance lines in the PAST tense (*es gab*, lit. 'it gave') are the most frequent with 80 tokens (40%). Ten percent of these are inflected for *Perfekt* (*hat gegeben*, lit. 'has given'). The second highest frequency, with approximately half as many tokens are sentences in pres_FUT, which, in turn, are equally frequent as utterances that are EXISTENTIAL or GENERIC. The second highest incidence is accounted for by the trinity of future intent (N=67 or 33.5%), pres_FUT, FUTURE, and IRREALIS. Alltogether, these amount to only 6.5 percentage points less than PAST tense, which makes the [*Regen+geben*] returns very future-oriented. Concordance lines annotated as pres_FUT encompass 27 modal verbs (as already noted in the last section, these are mostly instances of *kann* 'can', N=15). Strikingly, more than one-fourth of the future intent tokens are in the periphrastic FUTURE and almost half of these come from different types of literary texts, i.e., religious, song lyrics, fiction, and a proverb. The utterances in pres_FUT mostly belong to genre types WEATHER REPORT (N=19) and WEATHER TALK (N=16), where the latter mostly refer to weather forecasts, too. Only one pres_FUT utterrance comes from a LITERARY source. The lowest incidence of concordance lines involves PRESENT tense (N=15; 7.5%); more than half of these (N=9) are marked for PROGRESSIVE aspect, but these only occur in the PREC2 order.

Table 18. *Raw and relative frequencies of the TAM variable across the ARG+*geben*-Cx with the precipitate* Regen *'rain'*

Word order / TAM level	**PREC1** *Regen+geben*	**PREC2** *geben+Regen*	**Total**	**Relative frequency** /200
PAST	24	56	**80**	**40%**
PRESENT / prog	3/0	3/9	**15**	**7.5%**
pres_FUTURE	29	10	**39**	**19.5%**
FUTURE	18	0	**18**	**9%**
IRREALIS	8	2	**10**	**5%**
EXIS_GEN	18	20	**38**	**19%**
Total	**100**	**100**	**200**	**100%**
FUTURE_intent	55	12	67	33.5%
QUASI-AUXILIARIES	4	0	4	2%

Notably, the majority of future intent tokens occur in the PREC1 order, the majority of PAST tense tokens occur in the PREC2 order. EXISTENTIAL or GENERIC utterances are fairly evenly

distributed between the two word order constructions. The singular instance in *Perfekt*, exemplified in (63), is owing to *Regen* 'rain' being part of an anaphoric enumeration referring back to the antecedent pronoun *alles* 'everything'. Furthermore, the utterance rhymes (*Gewitter–bitter*), which indicates that the ME is part of a little poem. The stylized word order may explain its singular existence. Lastly, all concordance lines in the *Perfekt* tense belong to either genre type WEATHER REPORT (N=5) or LITERARY (N=3).

(63) ***Alles hat es hier bei uns gegeben,***
ob Wind, ***Regen****, Sonnenschein, Gewitter,*
das war nun ziemlich bitter! (heimat.eu 'home.eu')

'**It has given all here at our place**,
whether wind, **rain**, sunshine, thunderstorm
that was quite bitter.'

The concordance lines coded as WEATHER TALK genre predominantly attract PAST tense (67 out of 72 tokens, with no instances of *Perfekt* tense). Only four PAST utterances are part of the LITERARY genre and only a single concordance line is coded as WEATHER REPORT (as opposed to forecast), as in (64):

(64) *Im nördlichen Münchner Stadtgebiet gab es ein paar Blitzentladungen und* ***es gab*** *starken* ***Regen****.*
(sturmwetter.de – 'stormweather.de')
'In the northern Munich urban area there were a few lightning discharges and **it gave** heavy **rain**.'

With respect to quasi-auxiliaries, there are four tokens, all part of subordinate clauses in the PREC1 order, which are formed in a similar manner to the following example:

(65) *Die Wüsten bilden sich nicht,* ***weil es keinen Regen mehr gibt, sondern der Regen hört auf****, weil die Vegetation verschwunden ist.* (samen-des-wandels.de 'seeds-of-change')
'The deserts do not form because it does not give any more rain, but the rain stops because the vegetation has disappeared.'

Example (65) basically encompasses two quasi-auxiliaries: one is the "standard" *aufhören* 'stop' and the other indicates the stopping of the rain by using negation *kein* in combination with adverb *mehr* (*kein mehr* 'no more'). This type of negation is used in three out of the four utterances containing quasi-auxiliaries, in which the focus is more on the terminative than the incipient phase. In (65), the context expounds on the fact that the rain stopped indefinitely, rather than briefly. Lastly, there is one example where the start of the rain is indicated by the adverb *schon* 'already', i.e., *dass es schon Regen gibt* 'that it is already giving rain'.

3.1.2.3 The SPATIAL modifier variable

The distribution and frequencies of various SPATIAL modifiers for the [*Regen+geben*] construction are given in Table 19. The last column gives the relative frequency of the total per 200 concordance lines. The results sharply contrast with those of the [*Regen+fallen*] construction. Here, the vast majority of concordance lines (N=149; 74.5% of the data) are not modified by any SPATIAL expression whatsoever (ZERO). Where there is a SPATIAL modification, it tends to be a LOCATION with about one-fourth of the data (ten tokens provide the SETTING in the wider context). There are no instances of GOALS, SOURCES, or PATHS.

Table 19. *Raw and relative frequencies for the SPATIAL modification variable across the ARG+*geben*-Cx and the precipitate* Regen *'rain'*

Word order/ SPATIAL modifier	**PREC1** *Regen+geben*	**PREC2** *geben+Regen*	**Total**	**Relative frequency** /200
LOCATION	21	30	**51**	**25.5%**
ZERO	72	77	**149**	**74.5%**
GOAL	0	0	**0**	**0%**
SOURCE	0	0	**0**	**0%**
PATH	0	0	**0**	**0%**
Total	**100**	**100**	**200**	**100%**

In (66), a SOURCE and a GOAL-phrase modify the [*Regen+geben*] construction (*aus großen Wolken auf die Erde*) and possibly also a LOCATION (*da* 'there'). However, *da* could also denote temporal meaning 'then', which is not quite clear without further context. The second clause, which functions as a relative clause, contains a subject (*letzterer* 'the latter') rather than a relative pronoun and has its own verb (*sei gefallen* '[reportedly] fell')[77] whose path is modified by the SOURCE and the GOAL. Inarguably, the SPATIAL modifiers must be attributed to the [ARG-*fallen*] construction rather than [ARG-*geben*]. The combination of the two ARG-Cxx, [*Regen+geben*] and [*Regen+fallen*] (*Regen* is referred to as 'the latter'), nicely demonstrates how only one of the two constructions, i.e., ARG+*fallen,* easily incorporates SPATIAL modifications.

[77] *Sei gefallen* is *Konjunktiv I* ('subjunctive I' or 'special subjunctive'), which is used for reported speech. *Sei* is a modal verb and *gefallen* is the perfect inflection of *fallen* 'fall'. *Konjunktiv I* often expresses uncertainty, speculation, or doubt.

(66) ***Da habe es Donner, Blitz und Regen gegeben**, und **letzterer sei** offenbar **aus großen Wolken auf die Erde gefallen**, sagt der immer ein bisschen nach Streber riechende Wettermann*. (prisma.de)
'**There it (allegedly) gave thunder, lightning and rain**, and the latter apparently **fell from large clouds to the earth**, says the weatherman, who always smells a bit like a geek.'

With respect to the semantic nature of the LOCATION tokens, these can be grouped into four larger categories, which are "deictics", "location/particular place", "cardinal directions", and "other". "Deictics" make up the largest group and contain the most repetitive types, such as *hier* 'here' (N=4), *bei uns* 'here' (N=2); or both combined as *hier bei uns* (N=1), *da/dort* 'there' (N=3), and *örtlich* 'locally' (N=2). "Cardinal directions" include types like *im Norden* 'in the north' (N=3), *im äußersten Westen* 'at the extreme west', and also *in der Mitte* 'in the middle'. Furthermore, there are specific locations, mostly cities and countries, such as *in London*/*in Erfurt* 'in London/Erfurt', in *Spanien und Norwegen* 'in Spain and Norway', but also islands *auf Ibiza* 'on Ibizia', and certain places like *im Valley* 'in the valley'. There are not many combinations of LOCATIONS except for where a city is given as an antecedent, usually at the beginning of a text, sometimes in brackets, which is later referred to as *hier/bei uns* 'here'. There is one combination of a cardinal direction with a city: *im nördlichen Münchner Stadtgebiet* 'in the northern Munich city district'.

The [*Regen+geben*] construction does not readily lend itself to incorporate SPATIAL modifiers other than LOCATIONS. It is not that they are prohibited *per se*, but where they are included, the sentences sound odd at best with sources appearing more acceptable than goals (cf. ?/**es gab Regen auf die Erde* ?'it gave rain onto the earth'/?'there was rain onto the earth' vs. example (67) below). In contrast, LOCATIONS are unproblematic (e.g., *es gab Regen auf der Erde*/*auf Erden* 'it gave rain on earth'; the latter, *auf Erden*, is more biblical/tends to come from religious text sources).

(67) *Die Sonne war heute hier nicht zu sehen, **es gab Regen aus so gar nicht natürlichen Wolken** mit Rippen- und Wellenstrukturen*. (sevillana.de 'sevillana chemtrail diary')
'The sun was not visible here today, **it gave rain from not at all natural clouds** with rib and wave structures.'

There are two SPATIAL modifiers of PATH in the coded data frames at hand; however, these PATHS indicate the path of the human agents rather than that of the ME itself: *unterwegs auf der Rückfahrt zurück nach NRW gab es Regen* 'on the way back to NRW it gave rain' (cf. ex. (76), p.126) and *unterwegs gab es Regen* 'on the way gave it rain') and the ME is the secondary,

concomitant event.[78] Either way, the most salient result is the predominance of ZERO SPATIAL mentions in the returns for the [*Regen+geben*] construction. The next section deals with the NON-SPATIAL modifiers that accompany [*Regen+geben*] construction.

3.1.2.4 NON-SPATIAL modifier variable (TEMPORAL, QUANTITY, QUALITY, STANCE)

The distribution and frequencies of various NON-SPATIAL modifiers for the [*Regen+geben*] construction are given in Table 20. The last column shows the relative frequency of the total per 200 concordance lines. In total there are 347 modifiers (including SPATIAL modifications) in 200 concordance lines. In 123 concordance lines (61.5% of the data), there are at least two modifiers of a different type present. For the most part, I treat them independently.

Table 20. *Raw and relative frequencies of the NON-SPATIAL modifier variable across the ARG+*geben*-Cx and the precipitate* Regen *'rain'*

Word order / NON-SPATIAL modifier	**PREC1** *Regen+geben*	**PREC2** *geben+Regen*	**Total**	**Relative frequency** /200
TEMPORAL	52	41	**93**	**46.5%**
QUANTITY	47	51	**98**	**49%**
QUALITY	17	13	**30**	**15%**
STANCE	31	39	**70**	**35%**
Total	**147**	**144**	**291**	**146.5%**
ZERO	13	11	24	0.12%
COMBINATIONS	62	61	123	61.5%
SPATIAL modifier	27	20	47	23.5%

Regarding the total count of the data, there are 24 concordance lines where the weather event is not modified in any way (ZERO, 12% of the data). The most frequent NON-SPATIAL modifier marks QUANTITY and appears in almost half of the data closely followed by TEMPORAL modifiers with only five tokens less. The third most frequent modifier, by a small margin, is STANCE, which occurs in a little more than one-third of the data. Modifiers of QUALITY or MANNER are the least represented in the [*Regen+geben*] returns. Between the two word orders, the frequencies are distributed surprisingly evenly. This is also true for utterances which contain combinations of at least two modifiers or which contain ZERO modifiers. The most frequent stand-alone modifiers are TEMPORAL (N=24 in total; 12%) and QUANTITY (N=17 in total; 8.5%). The following subsections deal with each type of NON-SPATIAL modifier one by one.

[78] *NRW* is short for the western German state of *Nordrhein Westfalen* ('North Rhine-Westphalia).

Modifiers of TIME. As with the [*Regen+fallen*] construction, the high quantity of TEMPORAL modification brings with it a relatively high number of categories and types as well as some overlap. In general, there are the eight larger semantic categories of "day of the week", "time of day", "seasons", "frequency", "deictics", "duration/time period", and "other" or "combinations", which are listed by number of tokens per category in Table 21. For each category, some examples are given. The most frequent categories are "deictics" and "frequency". "Deictics" contains the most types overall as well as the most types which have more than one token; these types are listed in Table 21. Regarding "frequency", the tokens with the highest incidence are *zeitweise* 'occasionally/temporarily' (N=4) and variations of *wieder* 'again' (N=5). "Time of day", "months/seasons", and "other/combinations" are similarly frequent, however, here, too, there is overlap. For example, while the [*Regen+geben*] returns contain a decent number of concordance lines specifying a season (N=10), there are only two examples which make mention of specific months, which furthermore overlap with categories "deictics" as well as "duration/time period". Two "time of day" items include the day of the week (*in der Nacht zu Samstag* 'Friday night') and only one "day of the week" modifier encompasses the day only and no other temporal modifier. The remainder appear with modifiers of frequency and deictics. The category "other/combinations" contains utterances which are mostly idiosyncratic, such as *als wir aufbrechen wollten* 'when we wanted to leave' and a few are combinations such as *Montag, 14. März: am Nachmittag* 'Monday, March 14th… in the afternoon' or *Pfingsten… meistens* 'Whitsun… most of the time'. These were more difficult to assign to the other categories. There is only one concordance line, which includes an explicit date, even though no year is given in the concordance line itself.

Table 21. *Categories and frequencies of* TEMPORAL *modifiers occurring with the returns for the* Regen+geben *construction*

Semantic categories of SPATIAL modifications	**Total**	**Examples**
Deictics	27	*heute* 'today' *morgen* 'tomorrow' *gestern* 'yesterday' *zurzeit*/*momentan* 'now' *diesmal* 'this time' *bald* 'soon' *weiterhin* 'continuing'
Frequency	19	*zeitweise* 'occasionally/temporarily' *wieder* 'again', *immer wieder* 'again and again', *schnell wieder* 'soon again'
Time of day	13	*am Nachmittag* 'in the afternoon' *in den frühen Morgen- und späten Abendstunden* 'in the early morning and late evening hours' *bei Sonnenuntergang* 'at sunset'
Months/Seasons	12	*im Sommer 'in the summer'* *in der Regenzeit 'in the rainy season'* *in diesem August* 'this August' *von März bis Ende Mai* 'from March until the end of May'
Other/Combinations	11	*als wir aufbrechen wollten* 'when we wanted to leave' *auf der weiteren Fahrt* 'during the further journey' *beim Start* 'at the start' *Montag, 14. März: am Nachmittag* 'Monday, March 14th… in the afternoon' *Pfingsten… meistens* 'Whitsun… most of the time' *in der Nacht zu Samstag* 'Friday night' lit. 'in the night (leading up) to Saturday'
Duration/Time period	8	*von der Grundsteinlegung bis zum Richtfest* 'from the laying of the foundation stone to the topping-out ceremony' *binnen zwei Tagen* 'within two days' *bis zum Abend* 'until the evening' *bei ihrem Sommerfest* 'at their summer party'.
Day of the week	6	*Sonntag* 'on Sunday'
TOTAL	96	

Overall, it was difficult to tease the temporal modifiers apart unless they were simply annotated for being a "combination" of different modifiers or were counted twice. For example, *in diesem August* 'this August' contains both deixis and a specific month, which was categorized as "month" rather than as a deictic element or a combination and *in der Regenzeit* 'in the rainy season' clearly indicates a timespan, too, but was categorized as "month/season". However, stand-alone deictic items and modifiers of frequency stand out with regard to their high incidence and easily formed the categories so that they can clearly be stated as the semantic categories preferred by the [*Regen+geben*] construction.

Modifiers of QUANTITY. Almost half of the returns for the [*Regen+geben*] construction (N=98; 49%) contain a QUANTITY modifier. There are three observations that stand out regarding QUANTITY modifiers. First of all, there are hardly any quantities given in specific numbers. There are only two concordance lines which specify the quantity of rain that came down: one is given as *35l Regen pro Quadratmeter* '35l of rain per square meter' and the other as *mit 54mm etwas zu wenig* 'with 54mm a little too little'. Everywhere else, the quantities are mostly referred to by adverbs such as *viel* 'a lot' and *wenig* 'little'. Secondly, the items referring to QUANTITY are remarkably uniform. While there is a high number of tokens, there is a comparatively small number of types, which made the data easy to categorize in contrast to the other two constructions with *Regen* 'rain'. The modifiers are given in Table 22 and are discussed below. Lastly, there are many concordance lines where the quantity is ZERO, i.e., there was no rain. The ZERO quantities are mostly spelled out as *keinen Regen* 'no rain', but also as *keinen (einzigen) Tropfen* 'not a (single) drop' or *nicht ein bisschen* 'not a bit of'. Where some other participant is mentioned, the negation is given as *weder … noch* 'neither… nor'. Furthermore, there are a few instances where a large quantity or the type of rain is negated as in *fast keinen* 'almost no', *keinen kräftigen Regen mehr* 'no more heavy rain'. Almost 20% of the returns (N=39) include some sort of negation, which warrants a separate category in Table 22, in which the types of quantifiers are separated into four categories: negation, small quantity, large quantity, and other. There are only two instances of the latter.

Table 22. *Frequencies of modifier types and tokens in the* Regen+geben *returns specifying the quantity of rain as categorized by magnitude or absence of rainfall*

Negation		**Small quantity**		**Large quantity**		**Other**		**Total**
keinen 'no'	28	*etwas*	11	*viel*	16	*35l pro Quadratmeter*	1	
keinen Tropfen 'not a drop'	3	'some'		'a lot'		*'35l per square meter'*		
weder …noch 'neither… nor'	3	*wenig* 'little'	12	*reichlich* 'plenty'	3	*genügend* 'enough'	1	
		kaum 'hardly'	4					
		ein bisschen 'a little/a bit'	2					
other	5			other	9			
Total	**39**		**29**		**28**		**2**	**98**

The majority of QUANTITY modifiers specify the absence of rain using determiner *keinen* 'no' and , as in (68); however, the five items listed as "other" technically are not negators as they refer to miniscule quantities of rain like *kaum* 'hardly any'. If these were added to the category of small

quantities, it would be slightly more frequent than the "large quantity" category. Otherwise, the categories "small quantity" and "large quantity" are distributed almost evenly.

(68) *Ansonsten spielte das Wetter mit – **es gab keinen Regen** und im Laufe des Vormittags kam sogar die Sonne zum Vorschein*. (triathlon-bremen.de)
'Otherwise, the weather played along – **it gave no rain** and the sun even came out in the course of the morning.'

The most frequent items for small quantities of rain are *etwas* 'a bit/some' and *wenig* 'a little'. The latter also includes a couple of instances of *zu wenig* 'too little', *weniger* 'less', and *deutlich weniger* 'significantly less'. *Kaum* translates to 'hardly any' and *ein bisschen* is another synonym of 'a little'. While the "large quantity" column has a high number of *viel* 'a lot' and only three tokens of *reichlich* 'plenty', it also contains the highest number of unique items (N=9) consolidated as "other". In a similar manner as *wenig* 'little', *viel* 'a lot' includes variants such as *sehr viel* 'very much/a lot', *so viel* 'so much', *unendlich viel* 'endless', and *zu viel* 'too much'. Some of the unique—and more colloquial—items are *massig* 'shedloads', the expression *und das nicht zu knapp* 'and plenty of it' and the reduplication *Regen, Regen und nochmals Regen* 'rain, rain and more rain'. "Other" items are *genügend* 'sufficient' and the quantity in liters per square meter.

While the high total number of QUANTITY modifiers is intriguing in itself, it is also interesting that negation accounts for the highest frequency overall and that the unique items almost exclusively refer to larger quantities, with the exception of *genügend* 'sufficient'. At the same time, it seems straightforward that larger quantities are more salient and deemed worthy of reporting. We shall see that this is also true for the [*Schnee+geben*] construction, including the high incidence of negation, but neither for the [*Schneeregen+geben*] construction, which has a low incidence of quantitative modification but mostly mentions smaller quantities, nor for the [*Hagel+geben*] construction, which does include a lot of negation but varies otherwise.

Modifiers of QUALITY. Modifiers of frequency are somewhat infrequent (N= 30; 15%), but there is one type, which clearly stands out: there are eight tokens of *leicht* 'light', where one instance is furthermore defined as *anhaltend* 'lasting'. Furthermore, there is a second instance of *lang anhaltend* 'long lasting', and an additional two tokens each of *gefrierend* 'freezing', and *stark* 'heavy'. The remaining modifiers are unique types, such as *künstlich* 'artificial', *doof* 'stupid', *strömend* 'pouring', *heftig* 'severe', *prima* 'great', *sanft* 'soft' and *schauerartig*, *teils heftig* 'showery, sometimes heavy'. Other than the adjectival modifiers, there are three relative

clauses providing extra information on the noun phrase. These describe the rain as *einen Regen: der zum Teil auch kräftiger ausfällt* 'a rain which partly turns out heavier', *mit Geschmack und Nebelfratzenschneiden* 'with taste and fog making faces' (whatever that is supposed to mean), *einen Regen, der dir die Hosen ausgezogen hat* 'a rain that took your pants off' and lastly *Regen, deren Abflüsse die Leistungsfähigkeit der Kanalisationen überstieg, übersteigt bzw. übersteigen wird* 'rains, the discharge of which exceeded, exceeds or will exceed the capacity of the sewage system', as in the following example:

(69) *Fakt ist, **es gab, es gibt und es wird immer Regen geben**, deren Abflüsse die Leistungsfähigkeit der Kanalisationen überstieg, übersteigt bzw. übersteigen wird.* (institut-halbach.de)
'It is a fact that **there was, there is and there will always be rains**, the discharge of which exceeded, exceeds or will exceed the capacity of the sewage system.'

Last but not least, there are two nominal modifiers referring to a second participant, i.e., *ein Mix aus Regen und Sonne* 'a mix of rain and sun', as in (76) below, and *Regen mit Wind* 'rain with wind'.

Modifiers of STANCE. There is a comparatively high number of expressions of STANCE for these [*Regen+geben*] constructions with 70 out of 200 cases in total (35%). There are three categories of semantic prosody here: "positive" (N=26), "negative" (N=30), and "other" (N=16). One third of the data returns contain positive semantic prosody, which interacts with the high number of negated rain events, i.e., the speaker is hoping for no rain, as in (70), and is glad that it did not rain or that there is only little rain or only a short period of time where it rained. Only one utterance, by a runner, is positive about there being light rain (71). There is not a single utterance in the query returns with positive semantic prosody where rain is welcome after a dry spell, which is probably an indication of the fact that Germany usually does not suffer from droughts or forest fires, at least not to the extent found in southern Europe or California, to name just a few places.

(70) *Der Blick gen Himmel lässt **hoffen, dass es keinen Regen geben wird**.* (aktion-mensch.de - lit.'initiative human' – name of a German social organization)
'Looking at the sky gives **hope that it is going to give no rain**.'

(71) *Dabei waren die **Bedingungen** heute nahezu **traumhaft**, denn die Temperaturen lagen bei sage und schreibe 10 °C und **es gab leichten Regen, wunderbar***. (runningfreaks.de)
'The **conditions** today were almost **dreamlike**, because the temperatures were a whopping 10 °C and **it gave light rain, terrific**.'

Regarding utterances containing negative semantic prosody, almost all (28 out of 30) bemoan the fact that there was too much rain or that it will continue to rain. Examples of each are given in (72) and (73), respectively. There is one example about a people who sacrificed to their god when there was not enough rain. In (74), it is not quite clear whether the author is being ironic or not, as the full source text is no longer available. It appears to be meant ironically, however, if the speaker happens to be a farmer, it may well be a genuine complaint.

(72) ***Blöderweise*** *war es am Urlaubsort ziemlich schattig, und* ***es gab sehr viel Regen***. (die-klevers.de 'the-kleverfamily')
'**Unfortunately** (lit. '**stupidly'**), it was pretty shady at the resort and **it gave a lot of rain**.'

(73) ***Leider war das Wetter nicht gerade nett, es gab reichlich Regen*** *und Wind.* (keinsprint.de 'nosprint')
'**Unfortunately, the weather wasn't exactly nice, it gave plenty of rain** and wind.'

(74) *Und dann* ***ist*** *auch* ***der Sommer so schlecht****:* ***es hat noch keinen Regen gegeben***. (gutenberg.de)
'And then **the summer is so bad: it has not yet given any rain**.'

Lastly, there are 16 items, which are not positive or negative *per se*, but express surprise, doubt, or eventuality. Alternatively, the speaker makes some other observation about the weather, which is neither clearly positive or negative, as in (76), where the author refers to the weather as *seltsam* 'strange' and the sky as *interessant* 'interesting'.

(75) *Auch hier wurde es dunkel,* ***es gab Wolken und sogar Regen***. (psychoid.de)
'Here too it got dark and **it gave clouds and even rain**.

(76) *Heute auf der Rückfahrt aus der Eifel zurück nach NRW war* ***seltsames Wetter – es gab ein*** [sic] ***Mix aus Regen und Sonne – Interessanter Himmel.*** (bjklog.blogg.de)
'Today on the way back from the Eifel back to NRW **the weather was strange – it gave a mix of rain and sun - interesting sky**.'

In a similar manner to the [*Regen+fallen*] construction, STANCE is not expressed in an intense manner for [*Regen+geben*] construction. For negative prosody, *leider* 'unfortunately', *blöderweise* 'stupidly' and *doofer Regen* 'stupid rain' are the most emphatic collocating subjective comments about the weather situation. *Traumhaft* 'dreamlike' and *wunderbar* 'terrific', in (71), are the most positive remarks. In addition to *seltsames Wetter* 'weird weather', there is mirative *sogar* 'even', as shown in (75), which expresses the fact that rain was supposedly not expected, and one expression of *ich bin gespannt, ob* 'I am curious/excited if'. These are the most expressive assessments, which are neither positive nor negative but convey mild surprise at best. Most frequently, STANCE interacts with QUANTITY (40 cases in total). Unsurprisingly, these concordance lines contain many instances where negation and positive

semantic prosody interact (the speaker is glad about no rain) and instances where large(r) quantities interact with negative semantic prosody. While it makes sense for speakers to express their attitude with respect to certain quantities of rain and/or its absence, it is noticeable that modifiers of TIME, which are basically as frequent as expressions of STANCE, do not combine as readily with STANCE. In conclusion, four of the last five examples in this subsection contain a co-ME, which is discussed further in the next and final subsection for the survey of the [*Regen+geben*] construction.

3.1.2.5 The co-ME (co-occuring meteorological event) variable

Regarding the variable co-MEs, slightly more than half of the data, contain a co-ME. The distribution and frequencies are given in Table 23. The total is N=205 since in the PREC2 word order five multiple MEs and other MEs in the previous and/or following context are combined. The PREC2 word order contains roughly two-thirds as many co-MEs as the PREC1 order. The label "Multiple MEs" encompasses comma separated lists and coordinated MEs where *Regen* 'rain' is either the first, second, third, or fourth ME of several MEs, as in (66) and (75). The most frequent nominal ME collocates of *Regen* 'rain' in the *deTenTen13* are listed as *Wind* 'wind' (23,579 tokens and a *LogDice* score of 10.9), *Schnee* 'snow' (16,295 tokens and a *LogDice* of 10.7) and *Sonne* 'sun' (11,627 tokens but "only" a *logDice* of 9.7). *Kälte* 'cold' and *Sturm* 'storm' share the third highest *logDice* with 10.1, but only a frequency of 9,419 and 8,621 tokens respectively. The *logDice* indicates the collocational strength, which means that *Sonne* 'sun' more frequently collocates with *Regen* 'rain' than *Kälte* 'cold', but *Kälte* 'cold' and *Regen* 'rain' are a more typical collocation as shown by the higher *logDice* score (10.1 vs 9.7).

***Table 23**. Raw and relative frequencies of the co-ME variable across the ARG+*geben*-Cx and the precipitate* Regen *'rain'*

Word order/ co-ME	**PREC1** *Regen+geben*	**PREC2** *geben+Regen*	**Total**	**Relative Frequency** /200
Multiple MEs	21	35	**56**	**28%**
Previous clause	8	27	**35**	**17.5%**
Following clause	4	8	**12**	**6%**
Previous & following clause	1	5	**6**	**3%**
ZERO	66	30	**96**	**48%**
Total	**100**	**105**	**205**	**102.5%**

Mostly, co-MEs come in the shape of lists of several MEs (28% of the data): in most of the concordance lines (N=28), *Regen* 'rain' is the second ME listed (*es gab mildes Wetter und Regen*

'it gave mild weather and rain') and the first ME listed in another 23 utterances (*es wird Regen geben und Sonnenschein* 'it will give rain and sunshine'). Only in four cases is *Regen* 'rain' listed as the third ME and in one concordance as the fourth ME. A small number of these utterances contain zeugma, such as *Regen und Kurven* 'rain and curves' or *Regen und Schweine* 'rain and pigs'. An ME contained in the previous or following clause is often a noun as well, but it may also come in the shape of an adjective embedded in a different construction altogether (e.g., *es ist sehr windig und es gibt immer wieder Regen* 'it is very windy, and it gives rain again and again'; *die Luft war kalt und es gab Regen* 'the air was cold and it gave rain'). Obviously, the genre types WEATHER REPORT as well as WEATHER TALK that reproduces weather forecasts contain lists of MEs in either shape or form. In contrast to the [*es+fallen*] construction, it is easy to list different types of MEs with the predicate *geben* 'give'. This is especially true for the above mentioned MEs *Wind* 'wind', *Sturm* 'storm', *Sonne* 'sun', *Kälte* 'cold', which generally do not 'fall' but readily combine with *geben* 'give'. This subsection concludes the analysis of the [*Regen+geben*] construction and, thus, the two ARG-Cxx in combination with the rain precipitate. The next section concerns the *es*+PRED-Cx [*es+regnen*].

3.1.3 The predicate construction [es+regnen] *'it+rain'*

This section deals with the *es*+PRED-Cx [*es+regnen*] 'it rain'. The queries in the *deTenTen13* were carried out using the search strings in (i) and (ii). A total of 228,177 results were returned for both queries. Query (i) for the PREC2 word order returned more than twice as many results as query (ii) for the PREC1 word order.

(vii) [word="es"] [ ] {0,8}[lemma="regnen"] within <s/> (N=162,520)
(viii) [lemma="regnen""] [ ] {0,8}[word="es"] within <s/> (N=65,657)

For query (viii), 24 returns had to be discarded as false returns, due to the incorrect word order. All 24 turned out to be cases with a second event or activity containing the third person singular pronoun *es* 'it' as well. For query (vii), only one result turned out to be a spurious return in terms of word order. The following sub-sections deal with each variable and its corresponding sub-variables or factors, by giving in turn their raw (absolute) and relative frequencies as well as specific information about their distributions.

3.1.3.1 USAGE (LITERAL VS FIGURATIVE) AND GENRE VARIABLES

Table 24 gives the frequencies and distribution of the literal and figurative uses for the *es*+PRED-Cx and precipitate *Regen* 'rain'. The *es*+PRED-Cx is used predominantly in the literal sense in 192 of the 200 concordance lines returned. Each of the two word orders returned four instances of figurative use. The figurative cases turned out to involve transitive uses of the predicate, where it rains "direct objects" that have nothing to do with actual weather events, such as penalties, roses, criticism, gold and money. In the upcoming tables for each variable, these are categorized as "other" or "FIG" unless otherwise mentioned.

Table 24. *Frequencies of literal and figurative usage for the* es*+PRED-Cx and the precipitate* Regen *'rain'*

Word order / USAGE	**PREC1** *regnen+es*	**PREC2** *es+regnen*	**Total**
Literal	96	96	**192**
Figurative	4	4	**8**
Total	**100**	**100**	**200**

The frequencies for the GENRE variable as well as the distribution by word order are given in Table 25. The last column shows the relative frequency of the total per 200 concordance lines.

Table 25. *Distribution and frequencies of the GENRE variable across the* es*+PRED-Cx and the precipitate* Regen *'rain'*

Word order/ GENRE	**PREC2** *es+regnen*	**PREC1** *regnen+es*	**Total**	**Relative frequency** / 200*
WEATHER TALK	84	76	**160**	**80%**
WEATHER REPORTS	0	2	**2**	**1%**
METEOROLOGY	5	13	**18**	**9%**
LITERARY	7	5	**12**	**6%**
OTHER/FIGURATIVE	4	4	**8**	**4%**
Total	**100**	**100**	**200**	**100%**

*Total per level of GENRE divided by total of concordance lines for both word orders

Regarding the total frequencies, the vast majority of concordance lines is annotated as WEATHER TALK. These are fairly evenly distributed between the two word orders. The second most frequent type of genre, METEOROLOGY, only amounts to a little more than one-tenth of the WEATHER TALK data sets. Tokens from the LITERARY genre are even less frequent but fairly evenly distributed between the two word orders. Concordance lines that are figurative (FIG) or "other" account for eight concordance lines or 4% of the data. The least frequent, almost nonexistent with only two tokens in total, are WEATHER REPORTS. Since the corpus *deTenTen13* was chosen as a proxy for spoken language, the high number of WEATHER TALK tokens was somewhat expected.

However, the corpus also contains several newspaper sources, which could have increased the WEATHER REPORT count. We already saw in earlier subsections that the WEATHER REPORT type genre is more prone to attract the two ARG-Cxx rather than the *es*+PRED-Cx. The next subsection discusses TAM marking.

3.1.3.2 The TAM (TENSE, ASPECT, MOOD) variable

The absolute frequencies for the TAM variable as well as their distribution are displayed in Table 26. The last column gives the relative frequency of the total per 200 concordance lines. Regarding the total count, the majority of the events are talked about in the PAST tense. For the *es*+PRED-Cx, there is, indeed, some variety within the PAST tense factor: 19 concordance lines are nonfinite and another 16 are inflected for PERFEKT. The nonfinite concordance lines all combine with a (finite) quasi-auxiliary and focus on the incipient phase or the terminative phase of the event (*anfangen* 'start', as in (77), N=9; *beginnen* 'begin' and *aufhören* 'stop', each N=5, as in (80)), whereby the incipient phase is almost three times as frequent as the terminative phase. The individual frequencies of the quasi-auxiliaries are displayed in Table 27. Additionally, 11 concordance lines are annotated as progressive, as in (78), and two concordance lines occur in the PAST PERFECT.

(77) *Später wurden die Wolken dann bunt,* ***nach Einbruch der Dunkelheit begann es wieder zu regnen****, starker Wind kam auf.* (sevillana.de)
'Later the clouds became colourful, **after dark it started to rain again**, strong winds came up.'

(78) ***Mittlerweile regnete es ziemlich stark*** *und als wir in dem Dorf am See ankamen, blieben wir erstmal eine Weile im Auto sitzen*. (auswandererforum.de 'expatriatechatforum')
'**Meanwhile it was raining heavily** and when we arrived in the village by the lake, we stayed in the car for a while.'

Unsurprisingly, as the majority of the sample concordance lines belong to the WEATHER TALK genre, so too do they belong to the PAST tense. Otherwise, there are only three concordance lines, which are LITERARY, and one that is technically METEOROLOGICAL.

Table 26. *Raw and relative frequencies of the* TAM *variable within the* es*+PRED-Cx and the precipitate* Regen *'rain'*

Word order / TAM modifier	**PREC1** *regnen+es*	**PREC2** *es+regnen*	**Total**	**Relative frequency** / 200
PAST	32	43	75	37.5%
PRESENT / prog	8/30= 38	7/11= 18	56	28%
pres_FUT	2	2	4	2%
FUTURE	0	1	1	0.5%
IRREALIS	4	14	18	9%
EXIS_GEN	20	18	38	19%
FIGURATIVE	4	4	8	4%
Total	100	100	200	100%
Quasi-auxiliary plus non-finite PRED *regnen* 'rain'	6	28	34	17%

Table 27. *Distribution and frequencies of quasi-auxiliaries within returns of the* es*+PRED-Cx and the precipitate* Regen *'rain'*

Word order/ QUASI-AUXILIARY		**PREC1** *regnen+es*	**PREC2** *es+regnen*	**Total**
anfangen	'start'	0	12	**12**
beginnen	'begin'	1	10	**11**
aufhören	'stop'	1	6	**7**
other		4	0	**4**
Total		**6**	**28**	**34**

The second-most frequent TAM-marking category is PRESENT tense, with approximately 10 percentage points less than PAST. Approximately three quarter of the concordance lines coded as PRESENT convey progressive aspect. It is not always clear from the returns without further context whether an utterance is PROGRESSIVE or habitual, but where it is clear, it was coded appropriately. In comparison, EXISTENTIAL/GENERIC concordance lines, have the third highest frequency (N=38; 19%). The generic meaning of EXIS_GEN refers to the more habitual use. Examples (79) and (80) illustrate the difference between coding an utterance as PRESENT (79) or as EXIS_GEN (80): the former, through temporal modifier *im Moment* 'at the moment', imparts progressivity to the utterances. At the same time, *im Moment* does not mean (just) the very moment the speaker is producing the utterance, but, rather, a "weather routine" that seems to have been going on for a while, so that it appears almost habitual. While the weather situation in (79) may change at any time, this is not the case for (80) which describes how St. Peter makes it rain at any point in time.

(79) ***Also im Moment ist es so**, man steht frühmorgens auf und **es regnet** und es hat Nebel, heute hat es gestürmt wie im Herbst.* (obeabe.de)
'At the moment, it's like this, you get up early in the morning and **it rains** and there's fog (lit. it has fog), today, it has stormed as if it were autumn.'

(80) *…meinen deshalb, Petrus würde das Wetter machen, er würde mit seinem Schlüssel den Himmel aufschließen (**dann regnet es**) und wieder zuschließen (**dann hört es auf zu regnen**).* (predigtkasten.de sermonbox.')
'…think because of that St. Peter would make the weather, he would unlock heaven ('s gate) with his keys (**then it rains**) and lock it again (**then it stops to rain**).'

The three TAM factors with the lowest incidence are those, which can be subsumed under "FUT_intent", which taken together amount to 23 concordance lines or 11.5% of the data. The singular instance in FUTURE is furthermore negated (*Gott antwortet: es wird nicht regnen!!!* 'God answers: it is not going to rain!!!'). IRREALIS, which accounts for most of the FUT_intent tokens, mainly occurs in the PREC2 word order. Since IRREALIS instances contain a modal verb, the discrepancy in the distribution of the IRR annotation between the two word orders can mostly be explained by word order restrictions and style. Three of the four utterances in IRREALIS are negated (e.g., *hoffentlich regnet es nicht* 'hopefully it won't rain'). The 14 returns for the PREC2 order contain only one instance of negation, as shown in (81), and the majority (N=6) are variations of 'if/in case it rains' or 'should it rain', as in (82). The remainder are instances of *soll* 'supposed to', *kann/könnte* 'can/could' and *würde* 'would', all without overt negation marking.

(81) *Fragend schaute er zum Himmel. Es sah nicht aus, **als würde es noch einmal regnen**.* (homepagemodules.de)
'Questioningly, he looked up to the sky. It did not look **like it would rain again**.'

(82) *Und **wenn es regnen sollte**, wenn es schlechte Straßen sind, dann müssen sie ein Stück zu Fuß gehen.* (dkp-online.de *Die komplette Palette* 'The complete palette' – a socio-cultural art experiment)
'And **if it should rain**, if the roads are bad, then they will have to walk a bit.'

Otherwise, the types of TAM-marking are distributed somewhat evenly. The starkest discrepancy between *es+regnet* (PREC2) and *regnet+es* (PREC1) occurs in the PRESENT tense, where the PREC1 search query returned twice as many concordance lines in PRESENT or PRESENT PROGRESSIVE (more than three times as many instances are in present progressive) as PREC2. Six of the PREC2 order concordance lines that appeared in the PRESENT contain a quasi-auxiliary and focus on the start or end phase of the event (*beginnen* 'begin', N=3; *anfangen* 'start', N=2; *aufhören* 'stop', N=1). This is only true for two concordance lines for the other word order, which focus on the end phase (each one token of *aufhören* 'stop' and *nicht mehr* 'not anymore').

Take note that this only applies to the present tense instances. Overall, there are 34 utterances which contain a quasi-auxiliary, which is 18% of all literal concordance lines returned for the [*es*+*regnen*] queries. These instances clearly window the incipient phase of the event (N=25, including "other") over the terminative phase (N=9 including "other"). The MEs in these utterances usually coincide, often in a disruptive manner, with other events or activities the speaker engages in so that the moment rain sets in is salient. This becomes clear from what is often idiosyncratic modification further elaborated on in the following sub-sections.

In sum, most of the returned concordance lines are associated with WEATHER TALK and the majority were coded as PAST tense, followed by PRESENT tense. There is vast overlap between these two variables with 70 WEATHER TALK instances coded as PAST and 52 concordance lines coded as PRESENT. Obviously, we are prone to recount past events, especially if the weather played some role. The high amount of aspect marking (PROGRESSIVE) and quasi-auxiliaries furthermore illustrates the immediacy of the ME and the impact it has on our daily activities so much so that the moment that precipitation sets in or ceases is significant and worthy of being recounted and reported. This appears to be true in particular for a genre type like WEATHER TALK, which is supposedly more informal. A good deal of the WEATHER TALK returns are coded as EXIS_GEN and several of these instances follow subordinating conjunction *wenn* 'when/if', expressing a conditional clause as in an utterance where an allergic speaker states which weather situation is beneficial to them (*wenn es regnet, hat mein Leiden ein Ende* 'when it rains my suffering ends'). The returns with EXISTENTIAL or GENERIC use in utterances from the METEOROLOGY genre, as in (80) also tend to be introduced by conditional *wenn* 'when/if'. Only the LITERARY genre shows a wide variety of TAM factors despite its low frequency and attracts each TAM factor at least once and furthermore includes marking for *Perfekt*, progressive as well as quasi-auxiliaries. The eight figurative concordance lines are inflected for EXIS_GEN (N=2), PAST (N=2), PRESENT (N=2), PRESENT_progressive (N=1), and IRREALIS (N=1). The next section discusses SPATIAL modification of the *es*+PRED-Cx with *regnen* 'rain'.

3.1.3.3 The SPATIAL modifier variable

The returned concordance lines for the *es*+PRED-Cx with *regnen* 'rain' were annotated for five different types of SPATIAL modifications: LOCATION, SOURCE, GOAL, and PATH, as well as for absence of any SPATIAL mentions (ZERO). Table 28 shows the distribution of frequencies of different SPATIAL mentions occurring with the [*es*+*regnen*] construction. The last column gives the relative

frequency of the total per 192 literal concordance lines. Figurative meanings are not listed in the table; out of those eight only one contains a SPATIAL mention, a SOURCE, and the remaining seven do not encompass any SPATIAL mentions at all. The majority of literal utterances do not contain any SPATIAL modification whatsoever (N=105; 54.7%). The most frequent SPATIAL modification is a LOCATION. A total of 82 concordance lines (43%) denote some sort of specific place, where 54 (28%) directly modify the ME and another 28 tokens (15%) do not directly modify the rain event, but still provide a kind of SETTING; this is often the case for weather reports and meteorological descriptions of a region. The semantic nature of the SPATIAL modifiers is elaborated on below the table. There are only five concordance lines in total, which contain a SPATIAL modification that is not a LOCATION: two utterances contain a GOAL and three utterances contain a SOURCE. These five peripheral SPATIAL modifiers account for 2.6% of the data. There are zero PATHS and zero combinations of different SPATIAL modifications, such as a SOURCE and a GOAL .

Table 28. *Distribution and frequencies of SPATIAL modifiers across the* es*+PRED-Cx with the precipitate* Regen *'rain'*

Word order/	PREC1	PREC2	Total	Relative frequency
LOCATION	50	32	**82**	**43%**
GOAL	2	0	**2**	**1%**
SOURCE	2	1	**3**	**2%**
PATH	0	0	**0**	**0%**
ZERO	42	63	**105**	**55%**
Total	**96**	**96**	**192**	**101%**

With respect to the nature of the peripheral SPATIAL mentions, two of the SOURCES are very similar: *vom Himmel* 'from the sky' and *von dem dicht bedeckten Himmel* 'from the densely overcast sky'. The third one is more informal and idiosyncratic: *aus allen Rohren, die da oben geöffnet werden konnten* 'from every pipe that could be opened up there'. The two GOALS are more specific: *auf das wachsende Gras* 'onto the growing grass' and *ins Womo* (short for *Wohnmobil*, 'into the RV').

The LOCATIONS fall mostly into four categories semantically: "deictics", "cardinal direction", "location", "idiosyncratic", as well as "other". "Deictics" is the most frequent category with almost half of the modifiers (N=30), which consists almost exclusively the of adverbial types *draußen/draussen* 'outside' (N=14, the standard spelling is *draußen* 'outside', while the alternative spelling in Switzerland and Liechtenstein is *draussen* 'outside'. The spellings were kept as they originally occurred in the data) and *hier*/*bei uns* 'here' (N=11). Take

note that *SE's Word Sketch* function returns *draußen* 'outside' as the strongest collocation and one of the most frequent modifiers for the simple verb lemma *regnen* 'rain' within the *deTenTen13* corpus. *Bei uns* literally translates to 'with us' or 'at ours' and can mean 'at our house/place'; *hier bei uns* is also often combined and means something like 'over here' or 'here where we are'. The LOCATIONS occurring in the PREC2 word order almost exclusively consist of these adverbs, the exception being *im Umkreis von ca. 50km* 'within a radius of ca. 50km'. "Cardinal directions" are all unique types and are straightforward, such as *im Süden und Osten* 'in the south and the east' and *im Osten der Region* 'in the east of the region', but they can also be part of more complex combinations including a LOCATION, such as *an der Donau ca. 90KM westlich von uns* 'at the Danube about 90km west of us'. The "idiosyncratic" category overlaps with temporal items, e.g., *bei der Ankunft in Krabi* 'when we arrived in Krabi', *als wir dann in dem Großen Garten waren* 'when we then were in the Great Garden', *angekommen in Puerto Montt* 'when we arrived in Puerto Montt'. "Locations" can vary from simple places like cities (e.g., *in Berlin*), but also "zoomed in", very specific locations, such as *vorne bei der Käserei* 'at the cheese dairy ahead' and more complex combinations like *im Bereich der Rheinebene, vor allem in Mannheim und seiner direkten Umgebung* 'in the area of the Rhine plain, especially in Mannheim and its immediate surroundings', which intuitively appears to come either from a weather report or a meteorological description. The more idiosyncratic and specific locations appear chiefly in the PREC1 word order, which contains a higher frequency of LOCATIONS overall regarding both semantic categorization and syntactic parts-of-speech (hereafter PoS). The prevalence of LOCATIONS with PREC1 word order can be explained by word order rules. In a sentence that starts with an adverbial of place, the verb and subject must be inverted so that the verb remains in second position. The subject is then moved after the verb. It is slightly surprising that the [*es*+*regnen*] construction did not return more SPATIAL modifications in general (particularly if we ignored the LOCATIONS that do not immediately modify the ME but appear in the greater context) considering that FrameNet specifies [Place] —"the area which experiences precipitation"— as a core element of the precipitation frame.[79] Accordingly, we would expect an ME to specify the location where it takes place, possibly in the immediate context. This is clearly not the case. Neither is there a notable number of sources or goals. Since the [*es*+*regnen*] construction has presented itself as immediate and directly affecting the respective speaker's

[79] Source: framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Precipitation (last accessed December 6, 2012)

activities, it does not seem unreasonable to have expected more "experiencer goals", e.g., *es regnete mir ins Gesicht* 'it rained into my face' (lit. it rained me in the face'), where the speaker, as the experiencer of the ME, is possibly physically affected. The next section discusses the variable NON-SPATIAL modifiers.

3.1.3.4 The NON-SPATIAL modifier variable (TEMPORAL, QUANTITY, QUALITY, STANCE)

The returned concordance lines were also annotated for four different types of modifier: TIME, QUANTITY, QUALITY (and/or MANNER), and STANCE. Where there was more than one modifier of the same type (e.g., two specifications of time like the day of the week as well as the time of day), the concordance was only coded for the presence of a single modifier by type, but not how many. I furthermore annotated where there were no modifiers (ZERO, these instances contain no SPATIAL mentions either) and where there were combinations of two or more different types of modifiers (e.g., TEMPORAL+QUANTITY+STANCE). Lastly, where SPATIAL modification was present (e.g., a LOCATION, a GOAL, a SOURCE, or a PATH), these are also mentioned in Table 29 in the row "SPATIAL" and count towards the combinations of modifiers. The distribution and frequencies of modifiers for the *es*+PRED-Cx are given in Table 29. The last column gives the relative frequency of the total per 200 concordance lines (combinations are excluded from the total). Regarding the absolute frequencies, 20 concordance lines encompass no modifiers whatsoever (ZERO). The most frequent modifier by a wide margin is TEMPORAL, making up almost two-thirds of the returns. The second most frequent type of modifier are expressions of STANCE followed by modifiers of QUALITY, which occur less frequently than STANCE. The least frequent modification, by far, is that of QUANTITY. Even concordance lines that contain ZERO modification are slightly more frequent than modifications of QUANTITY. With respect to combinations, roughly two-thirds of the data contain at least two modifiers, SPATIAL and/or NON-SPATIAL. The most frequent combination of modifiers is that of TEMP+STANCE (N=26) and the most frequent stand-alone modifier is TEMPORAL (N=24) and then STANCE (N=15).

Table 29. *Raw and relative frequencies of the* NON-SPATIAL *modifier variable across the* es*+PRED-Cx and precipitate* Regen *'rain'*

Word order/	PREC1	PREC2	Total	Relative frequency
TEMPORAL	73	48	**121**	**61%**
QUANTITY	13	5	**18**	**9%**
QUALITY	38	21	**59**	**30%**
STANCE	33	47	**80**	**40%**
Total	**157**	**121**	**278**	**140%**
ZERO	2	18	20	10%
COMBINATIONS	77	50	127	64%
SPATIAL*	45	18	63	31%

*The numbers do not add up with the SPATIAL modifier variable in Section 3.1.3.3 since any SPATIAL mention in a return was additionally coded for but without teasing apart the types or considering combinations

The number of modifiers is higher for PREC1 for all types except STANCE. At the same time, the PREC2 order contains a much higher number of concordance lines with ZERO modification. The following sub-sections deal with each type of modifier one by in the order in which they are given in Table 29.

Modifiers of TIME. The temporal modifier, which is the most frequent modifier in the *es*+PRED-Cx, comes in various shapes and forms. As per *SE's Word Sketch*, five of the most frequent modifiers of the verb lemma *regnen* 'to rain' are temporal expressions including the most frequent one, which is *wieder* 'again'; furthermore, there are the deictic adverbs *heute* 'today', *gerade* 'right now', *morgen* 'tomorrow', and *gestern* 'yesterday'. Five of the ten most frequent modifiers in the *deTenTen1*3 are temporal and this fact is mirrored in the high incidence of temporal modifiers in the examined data. Often there are also two or more temporal modifiers combined. Regarding the linguistic properties of these modifiers, their syntactic and semantic use is detailed in the remainder of this section.

Basically, the section on the GENRE variable showed that the *es*+PRED-Cx almost never seems to be associated with WEATHER REPORTS (1%), while the vast majority of the returns constitute more informal WEATHER TALK (80%). This means that the TEMPORAL modifiers identified in the data at hand may very well differ from Liégeois's findings (2021), which examined weather reports prepared for different audiences. As noted above, it is not always easy to clearly categorize the temporal modifiers in the first place and they decidedly do not adhere to the classification used by Liégeois (ibid.), which is applied to the genre of weather reports only. His categories and findings are merely used as a comparison rather than as gospel and they may

come into play more for the ARG-Cxx, where there are more concordance lines coded as WEATHER REPORT.

Roughly, the modifiers of time can be grouped into seven semantic categories with at least five tokens each where the most frequent (N=25) is "deictics" and the most frequent token is *heute* 'today' (N=5). Other categories are "idiosyncratic", "frequency", "duration", "time of day", "months/seasons" and "combinations" as well as the catch-all category "other", which includes infrequent items such as sequence (*bevor* 'before') and day of week (am *Donnerstag* 'on Thursday'). There are 10 concordance lines, which include an exact date, but none where the date stands alone (at least not in the provided context). Another five tokens include the day and month, but not the year (*seit dem 15. Mai* 'since May 15th' and *ab dem 01. August abends* 'from the evening of August 1st'). However, the year may be given in the blog entry or message entry on the respective message board. The combined modifiers are time of day and the full date *in den frühen Morgenstunden am 21.06.07* 'in the early morning hours of 21.06.07'. There are only two simple "day of the week" tokens in the data, categorized as "other". Otherwise, days of the week are combined with a time of day, such as *am Donnerstagmorgen* 'on Thursday morning' or the more complex *nur am Samstag… in der Trainingszeit* 'only on Saturday… during practice' where the day of the week is specified as well as a duration.

The temporal element indispensable for a weather report is not necessarily required for a more informal WEATHER TALK situation. *Vice versa*, the often idiosyncratic temporal modifiers of a WEATHER TALK situation are neither appropriate nor likely to occur in a weather report. This is also due to the fact that these temporal modifiers tend to identify the often inopportune moment when rain sets in (e.g., *pünktlich zum Spielbeginn* 'just in time for the start of the game', *als wir kurz vor der Hütte sind* 'when we are near the hut') because it affects other activities/events, which are central to the respective speaker's story. In contrast, the central element of a weather report is the ME itself. However, speakers are inclined to use the activities they or others are engaged in to pinpoint exactly either the incipient or finishing phase of the concomitant weather event rather than supplying the time of day or the day of the week. An example is given in (83). Neither of the temporal modifiers (bolded) is likely to make an appearance in a more impersonal weather report.

(83) ***Exakt bei der Freigabe der Boxengasse zum 2. Qualifying*** *fing es wieder an zu regnen*.
(motorsport-xl.de)
'**Exactly when the pit lane was released for the second qualifying**, it started raining again.'

Other idiosyncratic examples of temporal modifiers include *als wir gerade wieder auf das kleine Boot umgestiegen sind* 'just when we got back on the small boat', *als die Kohle gerade begann zu glühen* 'as the coals just started to smoulder', or *seitdem ich das gesagt habe* 'since I've said that'. We have thus quite specific points in time that coincide with other activities. Similarly, one third of the temporal modifiers referring to the time of day mention a specific point in time (e.g., *um 15:15 Uhr* 'at 3:15pm'*, ½ 4* '3:30pm', *halb fünf* '4:30pm' lit. 'half five') rather than a more general part of the day like *Nachmittag* 'afternoon'. This specificity also applies to periods of time, such as *beim 58. Nationalen Sportfest des HSV …mehr oder weniger ununterbrochen* 'at the 58th national sports festival of the HSV .... more or less continuously', *in der zweiten Halbzeit* 'during the second half-time' or *die meiste Zeit während des Camps* 'most of the time during camp' as well as simply *dabei* 'at the same time' or *währenddessen* 'in the meantime'. The largest category, "deictics" is the clearest, in that there are repeat stand-alone instances of the same token, i.e., one type with several tokens. The already mentioned *heute* 'today' is not just the most frequent stand-alone modifier, but it occurs another six times with at least one other temporal modifier (e.g., *heute… gerade nicht...aber vor ca. 30 Minuten* 'today…not right now… but about 30 minutes ago', *11.12.2012 heute... immer wieder mal* '11.12.2012 today…again and again'). Furthermore, there are several progressive aspect-marking variations of 'now/right now' or 'at the moment', such as *nu*/*nun*, *zurzeit*, *jetzt*, *jetzt*…*eben* 'now…a moment ago', *jetzt*…*noch* 'now…still', *immer noch* 'still'. Consequently, adverbs, which greatly overlap with the semantic group of deictics, are the most frequent type of syntactic modifier with the constructions at hand.

Modifiers of QUANTITY. The smallest group of modifier type with only 18 tokens (9%) relates to expressions of QUANTITY. Quantity encompasses negation, as the quantity basically means zero. The modifiers can be grouped into four categories: "small quantity", "large quantity", "zero quantity" (negation), and "other" encompassing rate or comparison. The categories and their corresponding quantities and tokens are displayed in Table 30 below. There are five utterances in total, which contain negation (*nicht mehr* 'no more', two tokens; *gerade nicht* 'not right now', *hoffentlich nicht* 'hopefully not', and one quasi-auxiliary *aufgehört* 'stopped'). *Nicht mehr* 'no more/no longer' is the only type with more than one token. Another five utterances make mention of a small quantity by employing adverbs such as *kaum* 'hardly' or *wenig* 'little'.

Table 30. *Modifiers of* QUANTITY *across the* es*+PRED-Cx with the precipitate* Regen *'rain'*

QUANTITY	large	small	zero	other	Total
	zu viel 'too much' *teilweise* *echt viel* *sehr viel* 'very much'	*1. relativ wenig …* 'relatively little' *kaum* 'hardly' *ein bisschen* 'a little' *immer weniger* 'less and less'	*nicht mehr* 'no more' *nicht mehr* 'no more' *gerade nicht* 'not right now' *hoffentlich nicht* 'hopefully not'	2. *…600/650 mm Jahresdurchschnitt* '…600/650 mm yearly average' *35l pro m²* '35 liters per square meter' *etwa 30% mehr als in …* 'about 30% more than in …' *im Norden wesentlich mehr als im Süden* 'in the north significantly more than in the south'	
		nicht so viel wie von November bis April 'not as much as from November until April'	*hatte aufgehört* ('had stopped')	reduplication	
Total	**3**	**5**	**5**	**5**	**18**

In one case, the modifier is complemented by a subsequent sentence specifying the annual average in millimeters (cf. (84) and the items marked [1.] and [2.] in the table). This is a typical format for concordance lines from the METEOROLOGICAL genre, of which there are seven which contain a modifier of QUANTITY. Appropriately, these modifiers are accompanied by typical weather frame elements like a location, a time, or both, but never by a modifier of QUALITY. Modifiers of QUANTITY also occur with a comparison relating to location or time, e.g., *es regnet im Norden dieser Inseln wesentlich mehr als im Süden* 'it rains significantly more in the north of these islands than in the south' or *es regnet nicht so viel wie von November bis April* 'it does not rain as much as from November until April'. Tensewise, the concordance lines marked as METEOROLOGY coincide with TAM factor EXIS_GEN. This is also true for one of the utterances coded as WEATHER TALK. Otherwise, the utterances are mostly in the PAST tense except for one where the QUANTITY (but possibly also temporality) is expressed by reduplication (*Hier regnet und regnet und regnet es* 'It is raining and raining and raining here'). Three of the five negated concordance lines are part of the LITERARY genre, which itself is infrequent.

(84) *Im Bereich der Rheinebene, vor allem in Mannheim und seiner direkten Umgebung**, regnet es relativ wenig.** Die durchschnittliche Jahresniederschlagsmenge beträgt 600–650 mm.*
(naturefreunde-natura2000.de naturefriends-natura2000)[80]
'In the area of the Rhine plain, especially in Mannheim and its immediate surroundings, **it rains relatively little.** The average annual precipitation is 600-650 mm.

Considering the substantial majority of WEATHER TALK in the [*es*+*regnen*] data sets and the considerable lack of QUANTIFIERS within the very same, it appears that modifications of quantity are more likely to occur in other genres; the instances here indicate the METEOROLOGICAL genre and possibly also LITERARY discourse. The two instances of genre type WEATHER REPORT do not contain any QUANTITY modifiers and it remains to be seen whether the analysis of the other constructions and precipitates corroborate the observations made here.

Modifiers of QUALITY. There are 59 modifiers of QUALITY in total, representing about 30% of the returned *es*+PRED-Cx concordance lines involving rain. Semantically, the overwhelming majority of these modifiers describe the severity rather than the "lightness" of the ME. In fact, there are only two types and three tokens referring to light rain, i.e., *schwach* 'weakly' (N=1) and *leicht* 'lightly' (two tokens). For the "severity modifiers", the most frequent modifier is *in Strömen* 'it is pouring rain'/lit. 'in streams' (N=10) as well as variations of *stark* 'severely/heavily'/lit. 'strong' (N=9), such as *sehr stark* 'very severely', *ziemlich stark* 'quite severely', or *so stark, dass* 'so severely that'). Less frequent but more marked versions of *stark* include adverbs like *mächtig*/*kräftig*/*heftig* (all 'severely/hard') as well as the more informal *extrem* 'extremely', as in (85), *sintflutartig* 'torrentially', *wie verrückt* 'like crazy', and *wie sau* vulg. 'like crazy', which was also coded as a case of stancetaking. Other descriptions of severity are given in more unique analogies like *als würde der Himmel dafür bezahlt* 'as if the sky were paid for it', *nicht in Strömen, sondern in reißenden Flüssen* 'not in streams but in raging rivers', *aus allen Rohren, die da oben geöffnet werden konnten* 'from every pipe that could be opened up there'. Items that refer to the MANNER are *ununterbrochen* 'steadily' (N=4), as well as items that concomitantly express temporality (*noch kurz* 'briefly' and *lange* 'for a long time', two tokens). Last but not least, there is one more idiosyncratic utterance, which contains the co-ME

[80] *Natura 2000 is* a network of protected areas covering Europe's most valuable and threatened species and habitats. (Source: eea.europa.eu, accessed November 1, 2022)

Gewitterstürme 'thunderstorms', as in (86), as well as the judgment that this type of weather is "untypical" for the region of Silesia, Poland.

(85) *Hey ihr Lieben, könnt ihr mir vielleicht verraten, wo sich der Sommer versteckt hat? Es ist doch nicht mehr schön :(* ***Hier regnet es so extrem*** *und auch wenn es warm ist, das Wetter macht keine gute Laune*. (blogspot.co.at)
'Hey dears, can you maybe tell me where summer is hiding? It's not nice anymore :(**It's raining so extremely** here and even though it's warm, the weather doesn't put you in a good mood.'

(86) *Seit Wochen und* ***für Schlesien untypisch regnet es mit Gewitterstürmen*** (grafschaft-glatz.de 'county Glatz')
'For weeks and **atypical for Silesia, it has been raining with thunderstorms**.'

Regarding variable GENRE, modifiers of QUALITY occur twice within the METEOROLOGICAL genre and in both of the WEATHER REPORTS. The remainder appear solely within WEATHER TALK. With respect to combinations, a good number of QUALITY modifiers appear with (at least) a modifier of STANCE (N=19), as is true for (86). It makes sense for a qualitative statement to convey stance in itself (*wie verrückt* 'like crazy'; *wie sau*, vulg. 'like crazy') but also in the preceding and following context. More frequently, QUALITY modifiers are combined with TEMPORAL modifiers. However, as TEMPORAL modifiers occur within 61% of the PRED returns overall, this finding does not seem to be as remarkable. Most intriguing is the observation that there is not a single instance where a QUALITY and a QUANTITY modifier occur together no matter what the configuration. An explanation for this finding may lie in the similarity and possible overlap of QUANTITY and QUALITY. While I opted for QUANTITY to be encoded in lexical items denoting (literally) small and or large quantities, zero quantities, and "hard numbers" (cf. Table 30), a case could be made for items here coded as QUALITY (e.g., *wie verrückt/wie sau* 'like crazy' or *sintflutartig* 'torrentially') to concomitantly encode QUANTITY albeit in a less literal and more informal manner, which makes sense for WEATHER TALK items. That means, the established focus on "severity" and "force" as opposed to "lightness" could possibly also be interpreted as proxies for larger quantities. The reverse conclusion would be that light rain or little rain is lexicalized by more "formal" items (*kaum* 'hardly', *wenig* 'little') or possibly even completely different lexical items (e.g., *Sprühregen*/*Nieselregen* 'N. drizzle' or *nieseln* 'V. drizzle'). As well, the fact that a lot of rain or forceful rainfall is more salient overall and possibly disruptive to all sorts of activities may explain its predominance within the QUALITY tokens. Furthermore, just like QUANTITY and QUALITY appear similar to some extent, the items coded as QUALITY, again, share characteristics with modifiers of STANCE, which are surveyed in the next subsection.

Modifiers of STANCE. This section deals with attitudinal modifiers in which speakers convey their opinions and evaluations in relation to the expressed ME. The STANCE modifiers go hand-in-hand with semantic prosody. That is to say, ostensibly neutral words take on positive or negative connotations through frequent occurrences in certain semantic contexts or in collocations with specific items. However, here, it is also used for expressions that have rather obvious negative or positive associations. For example, the words *happen* and *cause* often occur in combination with negative outcomes, such as accidents (cf. Sinclair 1991, Stefanowitsch and Gries 2003). While rain is indispensable, it tends to impact mostly negatively all sorts of other events and actions that people engage in and which often take place outside. Accordingly, speakers take a negative stance when there is rain during holidays, when rain interrupts sports events (possibly even endangering participants as in racing events or cycling), or simply because in climes where there is no shortage of rain, it is associated with "bad" weather. Thus, many expressions of weather events, specifically in the *es*+PRED-Cx, are marked by negative semantic prosody and the corresponding modification.

A conspicuously high number of instances contain some sort of STANCE marker, i.e., 40% of the query returns. The PREC1 word order contains only about two-thirds of the amount of STANCE tokens that the PREC2 word order encompasses. Fourteen concordance lines contain only a STANCE marker and no other modifier, i.e., 17.5% of the *es*+PRED-Cx STANCE data. As the STANCE markers easily fall into four categories and are frequent enough, it was feasible to display the frequencies in a table. Table 31 shows the total counts and relative frequencies of modifiers as well as their frequencies based on whether they are positive or negative. The category "other" is comprised of STANCE expressions that are both positive *and* negative, in the very same utterance, or STANCE that is neither, such as surprise, irony, or resignation (the latter two, of course, often indicate a more negative attitude). Out of 80 concordance lines that express STANCE, 56 (70%) are negative. This tendency can also be gathered from SE's *Word Sketch* profile for verb *regnen* 'rain' since *leider* 'unfortunately' is listed as one of the most frequent modifiers of the verb lemma. In accordance with said list, *leider* 'unfortunately' turns out to be the most frequent modifier in the data set at hand as well (N=7; almost 9% of the returns).

Table 31. *Frequencies of the* STANCE *modifier and positive and negative* SEMANTIC PROSODY *across the* es*+PRED-Cx and the precipitate* Regen *'rain'*

Word order/			**Total**	**Relative frequency**	**out of***
Semantic prosody	**PREC1**	**PREC2**			
negative	20	36	**56**	70%	80
positive	10	7	**17**	21%	80
other	2	4	**6**	7%	80
out of STANCE total	32	47	79	40%	200

*80 is the total of concordance lines containing at least one STANCE token

Overall, negative stance seems to be expressed more emphatically in the examined concordance lines with with PREC2 order and through interjections such as *Mist!* 'crap!', *Bäh* and *bääääh* 'yuck/ew', as in (87), *übel* 'nasty', the vulgar *wie Sau* vulg. 'like crazy'/lit. 'like sow', and *was für ein mistiges Wetter* what a crappy weather'. Stance-like expressions with PREC1 constructions, as in (88), are more moderate. Seventeen concordance lines (or 21%) express positive semantic prosody, however, about half of these are negated in some shape or form. As for the positive stance in negated utterances, either the speaker is glad that it did not rain (e.g., *Gottseidank regnet es gerade nicht* 'thank God it is not raining right now'), hoping that it is not going to rain (*hoffentlich regnet es nicht* 'hopefully it is not going to rain'), or relieved that it stopped raining, as in (89), which is one of five concordance lines with a quasi-auxiliary.

(87) *Wettertechnisch schauˊs hier genauso aus...* ***bääääh!! Es hat die gaaaanze Nacht geregnet, nun könnte es mal wieder aufhören****.* (smilahome.blogspot.de)
'Weatherwise, it looks the same here…**eeeww**!! **It has rained aaaaaall night, it really could stop now**.'

(88) *Ist die Regenzeit Mitte Juli bis Mitte August an der Küste* ***tatsächlich so schlimm****. Regnet es wirklich von morgens bis abends? Ist an Baden zu denken??* (myanmar-guide.de)
'Is the wet season from mid of July until August **really that bad** at the coast? Does it really rain from morning to night? Is it possible to go swimming??'

(89) *Am Nachmittag* ***hörte es endlich auf zu regnen****, und das nutzten wir gnadenlos aus. (coppi-gym.de* 'Coppi high school')
'In the afternoon **it finally stopped to rain** and we took full advantage of that.'

Moreover, there are fourteen PREC2 utterances that feature the subordinating conjunctions *wenn*/*falls* ('if/in case') and all of them are coded for STANCE. Essentially, they express the possibility of rain and either the precautionary measures taken if it comes to the "worst", as in (90), or the unwelcome outcome, as in (91).

(90) *(…) und **falls es regnen sollte**, sitzen die Gäste gemütlich unter dem großen Vordach* (paz-online.de – 'online newspaper')
'(…) and **if it should rain**, the guests sit comfortably under the large canopy.'

(91) ***Wenn es regnen sollte**, wird sich das ganze Laub in eine schmierseifenähnliche Masse verwandeln und dann – Heidewitzka!* (gaertnerblog.de 'gardenerblog')
'**If it should rain**, all the leaves will turn into a soft soap-like mass and then – holy moly!'

Positive semantic prosody is found in many utterances by hobby gardeners, who lament a lack of rain, as in (92), allergy sufferers who experience relief once rain sets in (*mein Leiden hat ein Ende wenn* 'my suffering ends when'), or speakers for whom the rain is simply a welcome excuse to stay indoors and pursue activities like reading or gaming, as in (93). In the two examples below, the positive stance is explicitly communicated by the speakers (*ich bin richtig froh darüber* 'I'm really happy about it', *was gibt es da Schöneres* 'what could be nicer').

(92) ***Bei uns hat es jetzt** übrigens **auch eben zu regnen begonnen und ich bin richtig froh darüber**, es war alles schon so ausgetrocknet!* (blogspot.de)
'By the way, **it has just started to rain here and I am really happy about it**, everything was so dried out already.'

(93) *Wenn das kein Lesewetter ist! **Draußen regnet es**, es ist kalt und ungemütlich**, was gibt es da Schöneres,** als es sich mit einem guten Buch auf der Couch gemütlich zu machen.*
(toepfergemeinde.de 'potterycommunity')
'If this is not reading weather! **It's raining outside**, it's cold and uncomfortable, **what could be nicer** than curling up on the couch with a good book.'

Last but not least, one utterance refers to former German football legend and World Champion, Fritz Walter, who preferred to play in rainy weather. The turn of phrase *Fritz-Walter-Wetter* 'Fritz Walter Weather' is discussed in Chapter 4, Section 4.3.3.1.

3.1.3.5 *The co-MEs (CO-OCCURRING meteorological events) variable*

Lastly, regarding the variable that I call co-ME, more than half of the data, i.e., 56%, do not contain a co-ME. As a reminder, a co-MEs is any mention of another ME, which may occur, for example, as a list of precipitates or any other ME in a separate ME construction altogether. These may be different constructions or a combination of *es*+PRED-Cxx, such as *an dem Abend schneite und regnete es* 'that evening it snowed and rained'. The co-ME can be part of the KWIC, but it can also be part of the left or right context. Accordingly, it may appear in the clause preceding and/or following the target items. The frequencies are somewhat evenly distributed between the two word orders.

***Table 32**. Distribution and frequencies of the co-ME variable across the literal* es*+PRED-Cx and the precipitate* Regen *'rain'*

Word order/ co-ME	**PREC1** *Regen geben*	**PREC2** *geben Regen*	**Total**	**Relative Frequency** /192
ZERO	45	63	**108**	**56%**
PRED+PRED	4	8	**12**	**6%**
Previous clause	12	11	**23**	**12%**
Following clause	27	12	**39**	**20%**
Previous & following clause	7	2	**9**	**5%**
Other	1	0	**1**	**/**
Total	**96**	**96**	**192**	**0.99**

Co-occurring MEs in previous and/or following clauses are often found in coordinated clauses such as *es regnet und es gibt Wind/ein Gewitter* 'it is raining and it gives wind/a thunderstorm', i.e., they combine different MEs using different ME constructions like the *es*+PRED-Cx plus the *Regen*+*geben*-Cx just exemplified. As already mentioned, only a very small number of utterances combine MEs using the very same construction, i.e., the *es*+PRED-Cx, within the same clause. These PRED+PRED utterances may additionally express an ME in a previous or following sentence in a different construction. The one token of "other" contains a prepositional phrase in addition to the *es*+PRED-Cx: *es regnet mit Gewitterstürmen* 'it is raining with thunderstorms'.

This section concludes the survey of the *es*+PRED-Cx and its variables in combination with precipitate *Regen* 'rain'. The next section begins the analysis of the distribution and frequencies of the ARG-Cxx beginning with [*Regen*+*fallen*] construction. It compares the frequencies and distributions of the three constructions [*es*+*regnen*] (PRED), [*Regen*+*fallen*] (ARG), and [*Regen*+*geben*] (ARG) and elucidates the commonalities and differences between the usages of the three constructions.

*3.1.4 Distribution and frequencies of variables across the three constructions (ARG-*fallen*, and ARG-*geben*, es+PRED) with the precipitate* Regen *'rain'*

In this section, the frequencies and distribution of variables by construction, i.e., the *es*+PRED-Cx and the two ARG-Cxx, as surveyed up to this point, are compared against one another. We shall see that some variables appear to be constant and undeviating in their interaction with either construction, i.e., all three constructions predominantly associate with genre type WEATHER TALK, are mainly inflected for PAST tense, and frequently employ TEMPORAL modifiers. However, we shall also see that exploring these seemingly consistent variables reveals more subtle

differences, such as the diversity of TAM marking within the PAST tense instances, and the fact that some variables manifest considerable usage differences between the three constructions, such as encoding future events, stancetaking, and the use of spatial modifications.

3.1.4.1 *USAGE (LITERAL vs FIGURATIVE) and GENRE variables across the three constructions*

The distribution and frequencies of literal and figurative USE are given in Table 33. All of the constructions are overwhelmingly used in a literal sense. Only *Regen+fallen* (N=13) and *es+regnen* (N=6) contain somewhat low numbers of figurative instances. A total of 21 concordance lines are used figuratively; many of these are analogies. None of the second returns for the *Regen+geben*-Cxx contain any instances of figurative use.

Table 33. *Frequencies of literal and figurative USE by construction type (ARG+*fallen*, ARG+*geben*, and* es*+PRED) with the precipitate* Regen *'rain'*

USAGE / Construction	***Regen+fallen***	***Regen+geben***	**es+*regnen***	**Total**
LITERAL	187	200	192	**579**
FIGURATIVE	13	0	8	**21**
Total	**200**	**200**	**200**	**600**

The three constructions vary somewhat in terms of their use by GENRE as can be gathered from the raw and relative frequencies displayed in Table 34 and visualized in Figure 7. The absolute and relative frequencies show what has been established for the individual construction types: WEATHER TALK is the predominant type of genre in the *deTenTen13* corpus as well as in the returned data samples. Almost two-thirds of the returns are coded as WEATHER TALK. The second most frequent type of genre by a wide margin is METEOROLOGY, followed closely by utterances from the LITERARY genre. These utterances come from sources such as prose, religious texts, or more poetic language such as poetry or song lyrics. The least frequent type of genre is WEATHER REPORTS with 49 tokens (8% of the data). The remainder of "other/FIGURATIVE" amount to 3.5% of the data.

Table 34. *Raw and relative frequencies of the* GENRE *variable by construction type (ARG+*fallen*, ARG+*geben*,* es*+PRED) and the precipitate* Regen *'rain'*

Construction/ GENRE	*Regen+fallen*	*Regen+geben*	*es+regnen*	Total	Relative frequency
WEATHER TALK	87	130	160	**377**	**63%**
WEATHER REPORT	22	25	2	**49**	**8%**
METEOROLOGY	51	15	18	**84**	**14%**
LITERARY	26	30	12	**68**	**11%**
OTHER	1	0	0	**1**	**0%**
FIGURATIVE	13	0	8	**21**	**3.5%**
Total	**200**	**200**	**200**	**600**	**99.5%**

Pearson's chi-squared test: X-squared = 89.278, df = 8, p-value = 6.515e-16

The prevalence of the WEATHER TALK genre was foreseeable, considering the *deTenTen13* corpus was purposely sought out to ideally function as a proxy for spoken language. Yet, the low incidence of WEATHER REPORTS comes as a bit of a surprise as the *deTenTen13* appears to contain several online newspaper sources. Of course, the *deTenTen13* also contains weather reports not just by newspapers and their online presence, but also by privately owned blogs dedicated to reporting on the weather. These appear to be run most often by professional weatherpeople, but the language used reveals that the entries are written for a non-specialized audience, *aficionados* at best, who share their passion as weather enthusiasts, rather than by official and or public entities. These utterances are coded as WEATHER TALK rather than WEATHER REPORT.

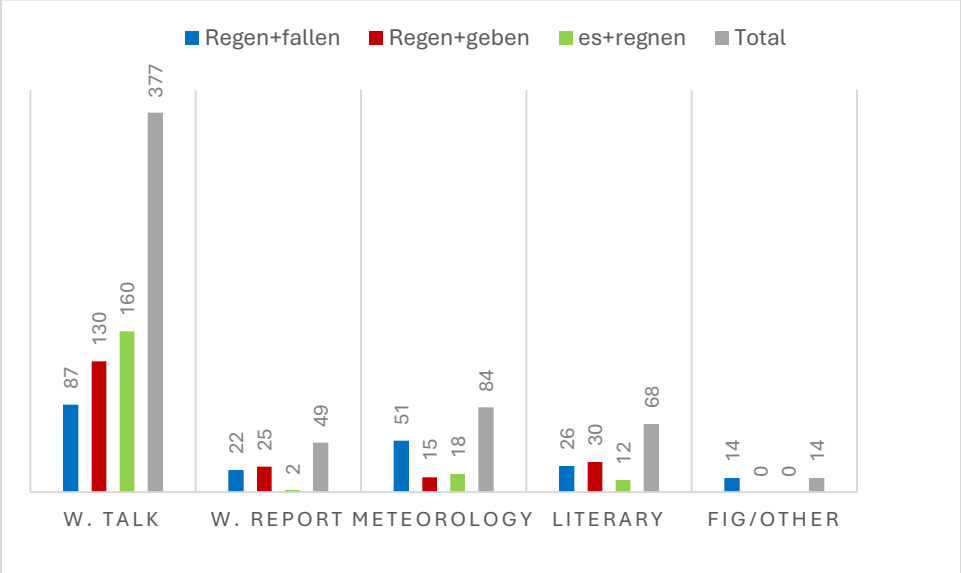

Figure 7. *Distribution and absolute frequencies of the three constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED-Cx) within the* Regen *'rain' precipitate data across the* GENRE *variable*

The stacked bar graph in Figure 8 shows how much each construction contributes proportionally to the absolute number of concordance lines per type of GENRE. It is obvious at first glance that

the data is heterogeneously distributed, i.e., each type of GENRE is prone to attract a different construction more than others. While it would seem intuitive for the two ARG-Cxx to coincide frequencywise–and not one of the ARG-Cxx and the *es*+PRED-Cx—this is not always the case. However, the ARG-Cxx only differ with six percentage points each with regard to genre types WEATHER REPORT and LITERARY. WEATHER TALK most frequently associates with the *es*+PRED-Cx, almost twice as much as with [*Regen+fallen*], possibly due to its more informal, or less specialized nature. In turn, the more formal, or specialized genre of WEATHER REPORT attracts both of the ARG-Cxx over the *es*+PRED-Cx.

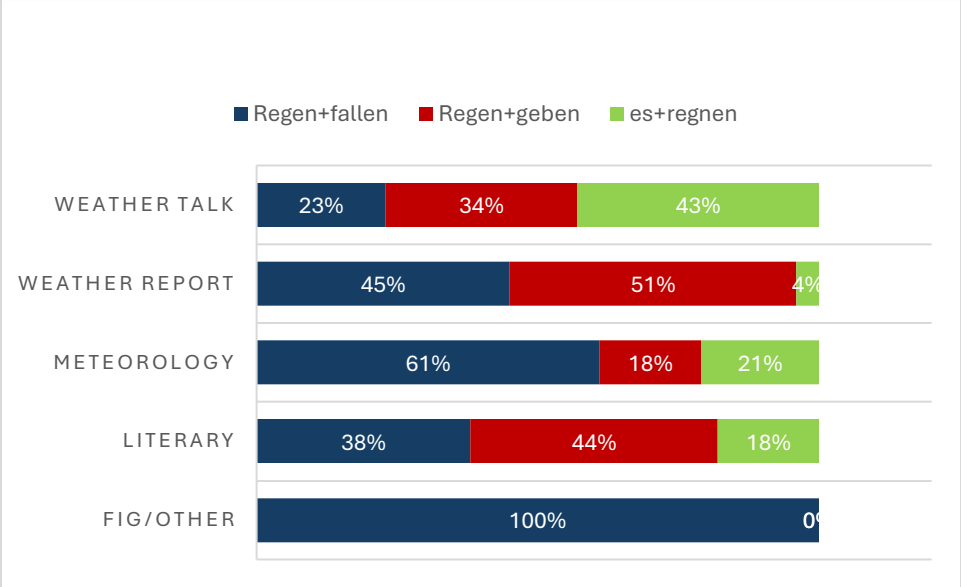

Figure 8. *Relative percentual frequencies of construction type (ARG+*fallen*, ARG+*geben*,* es*+PRED-Cx) with the precipitate* Regen *'rain' per type of GENRE*

The [*Regen+fallen*] construction sharply dominates the genre type METEOROLOGY and exclusively deals with HWW and AVW situations. In contrast, the returns for the *es*+PRED-Cxx are never used to explain "how weather works" but only to talk about general weather situations in specific regions/countries. The [*Regen+geben*] construction is a mix, which, in addition to HWW an AVW situations, contains more generic instances, such as *wenn es keinen Regen gibt, dann gibt es auch kein Wasser* 'when there is no rain, there is no water', in a text about cisterns. The LITERARY genre, like WEATHER REPORTS, predominantly attracts the two ARG-Cxx where five out of the 13 FIGURATIVE [*Regen+fallen*] returns also belong to genre type LITERARY. The [*Regen+geben*]

construction has the highest incidence of tokens that come from religious texts, as opposed to fiction or poetry, for example. This might have to do with an agentive, God, literally "giving" the people or the earth some type of precipitation, often to reward (with rain) or to punish (with torrential rains or destructive hail).

During the process of analyzing the data, the large variety within the WEATHER TALK genre became apparent with respect to content and the level of formality, which can be attributed to the fact that not all returns pertain to informal weather talk as a proxy for spoken language. If some of the sources were more easily and definitively traceable, it might have been worth it to tease apart the instances of WEATHER TALK which are *de facto* spoken versus those that were effectively prepared to appear as written text (e.g., reports, newspapers, articles, marketing) to allow a more comprehensive insight into those variables that operate more on the syntactic level. Such a step would not have influenced the frequencies of the other types of genre but may (or may not) have provided a more detailed picture of the genre type WEATHER TALK itself. To recapitulate, WEATHER TALK is the most frequent genre overall, which was expected due to choosing a corpus that was selected as a proxy for spoken language, but even though the three constructions are alike in predominantly containing WEATHER TALK, they clearly differ in the proportions of WEATHER TALK and consequently in the other types of genre they are attracted by. Finally, the vast majority of WEATHER TALK utterances are inflected for PAST tense, which is the overall most frequent tense, as noted in the next subsection, which deals with the TAM variable.

3.1.4.2 The TAM (TENSE, ASPECT, MOOD) variable across the three constructions

The distribution and frequencies of the TAM variable across the three constructions are given in Table 35; they do not include the tokens coded as FIGURATIVE or OTHER. The absolute frequencies are visualized in Figure 9 and the proportional percentage that each construction contributes to the total of each respective TAM factor are visualized in Figure 10. The absolute frequencies illustrate that speakers overwhelmingly report on rain events in the PAST; this is the case for more than one-third of the returns. But more so than discussing ongoing weather situations (PRESENT, roughly one-fifth of the returns) there is information out there about generic weather situations and discourse educating people on meteorology (EXIS_GEN, a quarter of the returns). However, rarely are ME constructions about rain inflected for FUTURE or IRREALIS.

Table 35. *Raw and relative frequencies of the* TAM *variable by construction type (ARG+*fallen*, ARG+*geben*, and* es*+PRED) and precipitate* Regen *'rain' (fig/other removed)*

Construction / TAM factor	*Regen+fallen*	*Regen+geben*	*es+regnen*	Total	Relative frequency /579
PAST	66	80	75	**221**	**38%**
PRESENT	34	15	56	**105**	**18%**
pres_FUTURE	18	39	4	**61**	**11%**
FUTURE	0	18	1	**19**	**3%**
IRREALIS	0	10	18	**28**	**5%**
EXIS_GEN	69	38	38	**145**	**25%**
Total	**187**	**200**	**192**	**579**	**100%**

Pearson's chi-squared test: X-squared = 117.69, df = 10, p-value < 2.2e-16

Figure 9 reveals that of all categories of TAM, the use of PAST is the most frequent overall, as well the most homogeneously distributed among the three constructions with roughly one-third of the data per construction. While PAST tense is the most frequently occurring tense for both the *es+regnen* and the [*Regen+geben*] returns, it is not the most frequent tense with respect to the [*Regen+fallen*] returns, which most frequently attract EXIS_GEN by a slim margin.

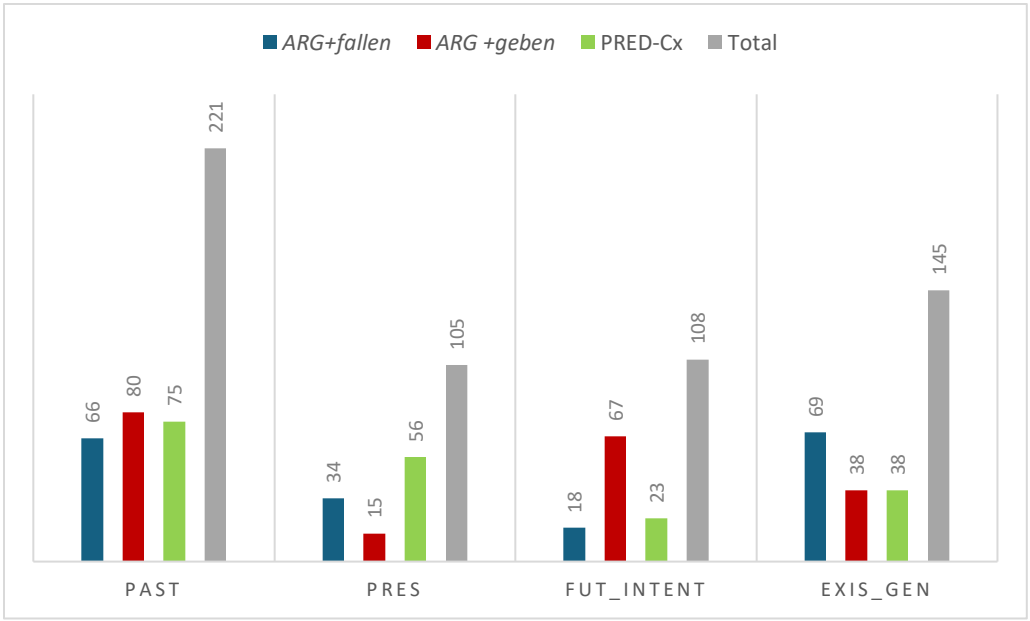

Figure 9. *Absolute frequencies and distribution of* TAM *factors across the three constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED-Cx) with the precipitate* Regen *'rain' and with categories pres_*FUT*,* FUT*, and* IRR *subsumed under* FUT*_intent ("future intent")*

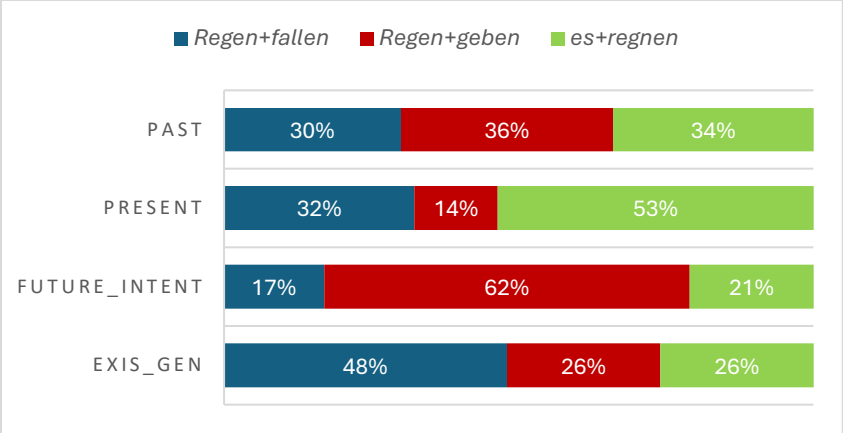

Figure 10. *Relative frequencies in percentage of each individual construction (ARG+*fallen*, ARG+*geben*, es+PRED) per total of TAM category within the precipitate* Regen *'rain' data sets, present_FUTURE, FUTURE, and IRREALIS are subsumed under FUTURE_intent*

The following subsections go into slightly more detail with respect to each TAM factor starting with PAST tokens and the differences within the PAST tense data at hand. The three constructions each attract variable PAST with roughly one-third of the returns, but at closer examination (aspectual) differences are revealed. The frequencies of the factors present progressive, *Perfekt* and past perfect, as well as quasi-auxiliaries across all three constructions are stated in Table 36.

Table 36. *Frequency and distribution of PAST tense differences across the three constructions (ARG+*fallen*, ARG+*geben*, and* es*+PRED) and the precipitate* Regen *'rain'*

Construction TENSE/ASPECT	***Regen+fallen***	***Regen+geben***	***es+regnen***	**Total**
Past progressive	2	0	11	**13**
Perfekt	1	8	16	**25**
Past perfect	0	0	2	**2**
Quasi-auxiliaries	2	4	18	**24**
Total	**5**	**12**	**47**	**64**
PAST **total**	66	80	75	221

The [*es+regnen*] construction is clearly the most diverse with regard to aspect marking and accounts for approximately three-fourth of the aspect marking, quasi-auxiliaries, and past perfect inflections overall. Almost two-thirds of the [*es+regnen*] returns coded as PAST contain some aspect marking (mostly past perfect) or a quasi-auxiliary, which focuses the attention on the incipient or terminative phase of the ME. In combination with especially the idiosyncratic TEMPORAL modifiers, which pinpoint very specific moments in time, these collocations bear witness to the impact the ME has on other concomitant activities or events but also indicate how the ME is not necessarily the focus of the discourse. In contrast, the ARG-Cxx barely contain

differences within PAST tense, *Regen+geben* more so than [*Regen+fallen*], which makes them less dynamic. To some extent, the "conceptual orality" of *Perfekt* tense is also given for the [*Regen+geben*] returns and to an even lesser extent the focus on the incipient or terminative phase of the ME can be found, too. While the *Regen+fallen* returns have the lowest incidence of PAST tokens overall as well as the lowest variety within the PAST data, they overwhelmingly convey existential or generic meaning. EXIS_GEN is not just the most frequent TAM category overall for the [*Regen+fallen*] returns, but [*Regen+fallen*] also accounts for almost half of the EXIS_GEN tokens across all three constructions. The preponderance of EXIS_GEN with the [*Regen+fallen*] construction is intriguing, considering that [*Regen+geben*] is considered to be *the* existential construction in German, which is typically translated as 'there is' in English. Mostly, these concordance lines interact with the METEOROLOGY genre and describe the general weather conditions in a specific region or place. While existential or generic use, of course, also relates habitual use, more than half of the returns coded as PRESENT tense (57%) carry progressive aspect. Again, the highest incidence, almost two-thirds, of progressive aspect is contained within the [*es+regnen*] returns, which corroborates the finding for PAST tense, that this construction exhibits a high level of immediacy. [*Regen+fallen*] attracts about one-third of the PRESENT tense tokens and one-fourth of all progressive aspect tokens and the [*Regen+geben*] construction contains roughly 15% of all PRESENT tense tokens and all progressives, likewise corroborating the findings from the PAST tense data.

The categories FUTURE, IRREALIS, and present_FUTURE, taken individually, are rarely used to talk about rain events. However, subsumed as "future intent" tokens, they are even slightly more frequent than PRESENT tense tokens. Strictly speaking, as none of these events have yet been actualized, they belong in the realm of irrealis, or more fittingly—particularly regarding weather forecasts—in the realm of *potentialis*. By this, I mean that there is a probability that the event will happen. In fact, weather forecasts are frequently presented with a percentage chance or likelihood of precipitation. The comparatively low incidence of WEATHER REPORTS, particularly for the [*es+regnen*] returns, is mirrored in the relatively infrequent proportion of concordance lines with "future intent" (14%), especially regarding future and irrealis marking. Having teased apart the frequencies reveals a clear trend for the *Regen* 'rain' precipitate: [*Regen+geben*] is not just the "designated" construction used for "future intent", but out of 19 FUTURE tokens, 18 are provided by the [*Regen+geben*] construction. Hence, FUT appears to be almost exclusively reserved for *Regen+geben* constructions, at least within the returns at hand. Both *Regen+fallen*

and *es+regnen* are infrequent, however, they differ in that several of the [*es+regnen*] returns contain modal verbs (N=22), which is not true for the [*Regen+fallen*] returns (N=5). [*es+regnen*] thus appears to be the construction least likely to lexicalize a prospective or potentialis meaning and instead contains more utterances where hopes and wishes are expressed. Considering that weather forecasts deal with predictions that can only be made somewhat reliably for three days or so in advance, it seems intuitive for genre WEATHER REPORT to attract irrealis mood and, in fact, [*Regen+geben*] tokens with IRREALIS overwhelmingly interact with this genre type (more than half of the tokens), so do the instances from the *Regen+fallen* returns. In comparison, future intent utterances from the [*es+regnen*] returns belong either to genre type WEATHER TALK or LITERARY.

In conclusion, it makes sense for PAST tense to be the most frequent level of variable tense across constructions since speakers tend to talk about things or events that "happened" in the past. Proportionally, PAST is distributed surprisingly homogeneously between the constructions, but this is where the similarities end. [*Regen+fallen*], exceptionally, attracts existential or generic meaning despite [*Regen+geben*] generally being the designated "existential construction" in German. By contrast, [*Regen+geben*] notably attracts utterances in the realm of irrealis or what I dubbed "future intent", while the [*es+regnen*] construction appears to be firmly grounded "in the present". It is highly immediate and focused on the direct experience of the ME and the affect it has on other, concomitant events and activities. Even in PAST tense, and unlike the ARG-Cxx, the *es*+PRED-Cx exhibits most diversity and presents a pot-pourri of *Perfekt* (an indicator of orality), past perfect, progressive aspect, and overall, a high number of quasi-auxiliaries focusing in particularly on the incipient phase of the rain event.

3.1.4.3 The SPATIAL modifier variable across the three constructions

The distribution and frequencies of SPATIAL modifiers across the three constructions are given in Table 37. The distribution and absolute frequencies are furthermore visualized in Figure 11. More than half of the returned concordance lines do not encompass any SPATIAL modification whatsoever (56%). Where there is SPATIAL modification, it is usually a LOCATION (37%). The second most frequent SPATIAL modifier is a GOAL (4%), followed by a SOURCE (2%). PATHS are the most infrequent SPATIAL modifiers and only present within the [ARG+*fallen*] returns (1%).

The most obvious finding is the predominance of ZERO SPATIAL modification; this is true for the [*Regen+geben*] construction in particular. In fact, there are more than twice as many [*Regen+geben*] concordance lines without a SPATIAL modifier than [*Regen+fallen*] concordance

lines without a SPATIAL modifier. The [*es+regnen*] construction is somewhere in the middle but still contains no SPATIAL modification in more than half of the returns. [*Regen+fallen*] is the only construction which has more instances of actual SPATIAL modification, i.e., LOCATIONS, than ZERO SPATIAL modification, if only by a small margin. LOCATIONS are almost evenly distributed between the [*es+regnen*] and the [*Regen+geben*] returns, however, the constructions differ in which type of SPATIAL modification they attract semantically: [*es+regnen*] returns predominantly contain two types of "deictics", i.e., *draußen*/*draussen* 'outside' and *hier*/*bei uns* 'here'. [*Regen+geben*] returns contain a larger number of different "deictics" types. In contrast, the [*Regen+fallen*] construction attracts a completely different set of SPATIAL modifiers, i.e., "goals", "sources", and the most frequent type of LOCATION is a "specific place" rather than "deictics".

Table 37. *Distribution and frequencies of the SPATIAL modification variable across constructions (ARG+*fallen*, ARG+*geben*, es+PRED) with the precipitate* Regen *'rain'*

Construction/ SPATIAL modifier	***Regen+fallen***	***Regen+geben***	***es+regnen***	**Total**	**Relative Frequency /579**
LOCATION	80	51	82	**213**	**37%**
GOAL	22	0	2	**24**	**4%**
SOURCE	11	0	3	**14**	**2%**
PATH	6	0	0	**6**	**1%**
ZERO	73	149	105	**327**	**56%**
Total	**192**	**200**	**192**	**584**	**100%**

Pearson's chi-squared test: X-squared = 184.39, df = 4, p-value < 2.2e-16.

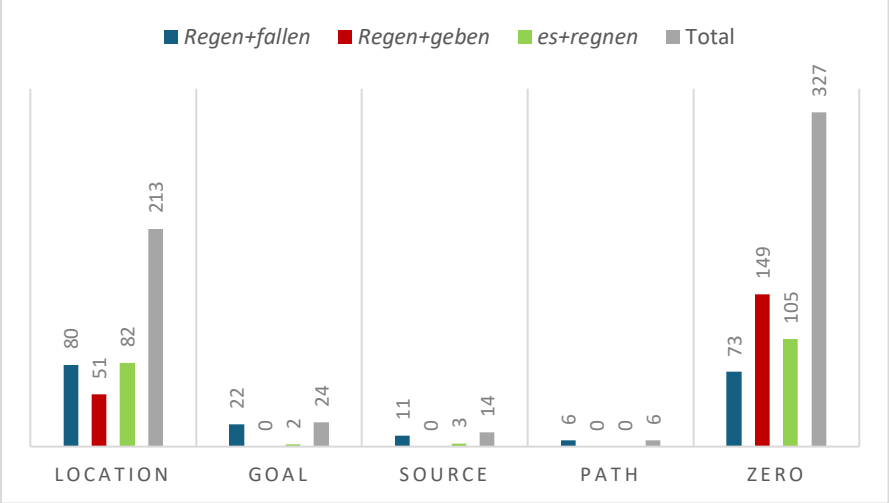

Figure 11*. Distribution and absolute frequencies of SPATIAL modifiers across the three constructions (ARG+*fallen*, ARG+*geben*, es+PRED) with the precipitate* Regen *'rain'*

It makes sense that the most frequent PATH, where there is one, is a LOCATION. Clearly, *where* the ME is happening is of fundamental importance. In that regard, it seems surprising that there are

many concordance lines seemingly without any SPATIAL mention. Obviously, the data are limited contextwise, so that a lot more SPATIAL mentions, particularly LOCATIONS, may be found in the even wider context. Going back to the frames established by FRAME-semantics, we would assume every single event to contain some kind of SPATIAL modification as it is a core element of both the WEATHER- and the PRECIPITATION frame. The fact that GOALS constitute the second most frequent SPATIAL modification, albeit by a wide margin, is plausible as well. We know that rain comes "from above", generally speaking, but the goal may vary, although it is ultimately, of course, "down" or onto the ground. Thus, where a location is given the goal is, in some way, implied. Where a goal is explicitly mentioned, it is either very general, as in HWW utterances (*zur Erde* 'to the ground/earth'), or very specific as in *auf die Stiefmütterchen* 'onto the pansies' or *aufs Vordach über ihnen* 'onto the roof above them'. The rain does clearly not fall only onto the specific goals, but 'the roof above them' is where the rain produces the pattering sound and, presumably, where the roof keeps 'them' dry. The mention of the pansies is probably due to the flowers either being in need of watering or because they are in danger of being destroyed by too forceful a rainfall.

In the extensive research on motion events, it has been established that there is a "goal bias" in many languages (Ikegami 1987, Lakusta and Landau 2005, Stefanowitsch and Rohde 2004). However, Verspoor and Dirven (2004: 84f.) point out that "For *human action*, the goal is usually more important than the source and the source and goal are usually more important than the path" (italics added by me). For non-volitional, inanimate agents like the weather, this does not seem to apply. Guse (2022), in a corpus study of the source-goal asymmetry in German, found that, surprisingly, volition does not influence the likelihood of windowing the goal path. However, in her data, goal paths were significantly more likely to be expressed with human agents when compared to animal agents. Having said that, non-animated subjects windowed the goal path the most.

In summary, [*Regen+fallen*] is the construction that contains the most SPATIAL modification and the most variety of SPATIAL modification while ARG-*geben* is the construction that is the least likely to attract any SPATIAL modifiers, and where there is one, it tends to be a LOCATION. The *es*+PRED-Cx is somewhere in the middle of the two ARG-Cxx with respect to ZERO SPATIAL modification and LOCATIONS. However, regarding the less frequent SPATIAL modifiers, GOAL, SOURCE, and PATH, it is closer to [ARG-*geben*] in terms of frequency, but, nevertheless, it does encompass a handful of GOALS and SOURCES.

3.1.4.4 The NON-SPATIAL modifier variable across the three constructions

The distribution and frequencies of NON-SPATIAL modification across the three constructions are given in Table 38 (the count of SPATIAL modifiers is added in the last row), and are visualized in Figure 12. The proportional percentage that each construction contributes to the total of the respective modifier is shown in Figure 13. Regarding the total frequencies, it is intriguing how similar the total frequency of collocating modifiers is per construction even if the modifiers themselves are distributed somewhat differentially. Only 3% of the data overall do not contain any type of modifier, NON-SPATIAL or SPATIAL. Despite having the lowest frequency overall of modifiers, the [*es+regnen*] construction returns also contains the lowest number of ZERO-modifier constructions. The highest number of concordance lines without any type of modifiers pertains to *Regen+geben*. Overall, about one-third of the concordance lines encompass at least two modifiers, NON-SPATIAL and/or SPATIAL where, again, the ARG-Cxx are closer in number while the [*es+regnen*] construction returns contain the lowest number of combined-modifier constructions.

Table 38. *Instances and frequencies of NON-SPATIAL modifier types across the three constructions (ARG+*fallen*, ARG+*geben*, es+PRED) with the precipitate* Regen *'rain'*

Construction/ NON-SPATIAL modifier	***Regen+fallen***	***Regen+geben***	***es+regnen***	**Total**	**Relative frequency /579**
TEMPORAL	124	93	121	**338**	**58%**
QUANTITY	91	98	18	**207**	**36%**
QUALITY	46	30	59	**135**	**23%**
STANCE	26	70	80	**176**	**30%**
Total	**287**	**291**	**278**	**856**	**147%**
COMBINATIONS*	138	123	77	338	(/1047) = 32%
ZERO*	9	24	2	35	(/1047) = 3%
SPATIAL modifier	82	47	62	191	/

*The rows COMBINATIONS and ZERO take SPATIAL modifiers into account, i.e., the relative frequency is divided by 1,047 (856 plus 191)

Pearson's chi-squared test: X-squared = 100.02, df = 6, p-value < 2.2e-16

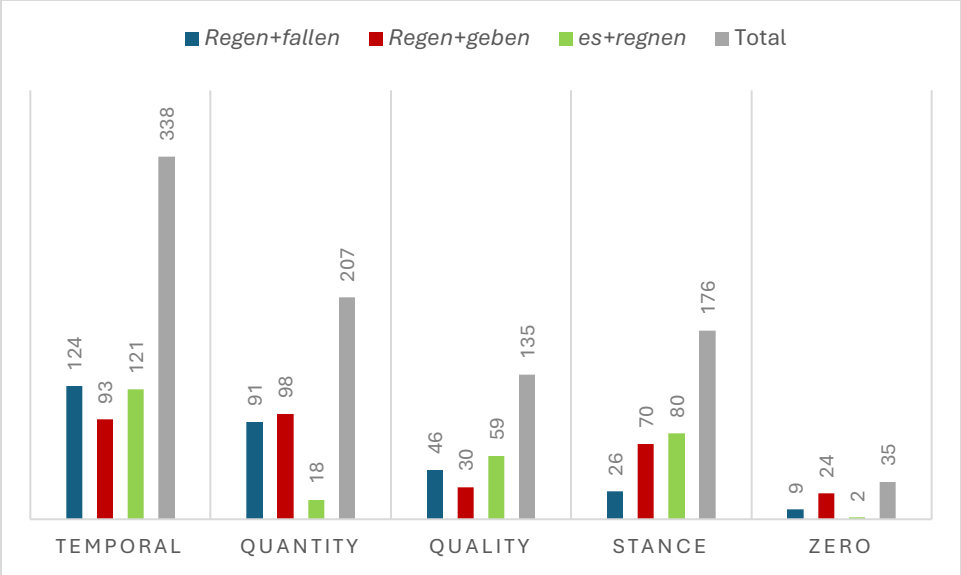

Figure 12. *Absolute frequencies of* NON-SPATIAL *modifiers across the three constructions (ARG+*fallen*, ARG+geben,* es*+PRED) with the precipitate* Regen *'rain'*

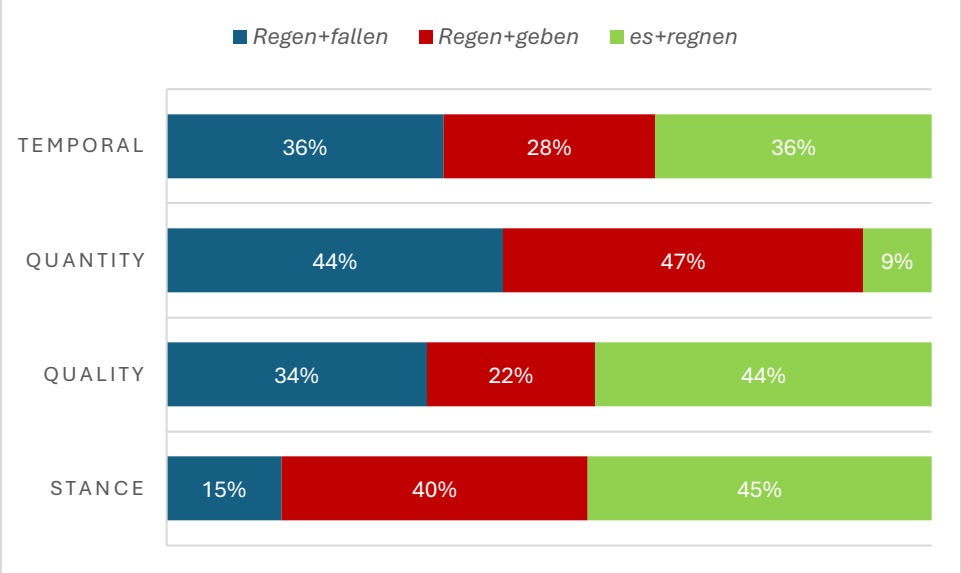

Figure 13. *Percentual proportion of each individual construction (ARG+fallen, ARG+geben,* es*+PRED) per total of* NON-SPATIAL *modifier within the precipitate Regen 'rain' data sets*

Figure 12 sheds light on the distribution of the NON-SPATIAL modifier variable across the three macro constructions. It does not include COMBINATIONS and/or ZERO modifiers. Figure 12 shows a generally heterogeneous distribution of modifiers and constructions with respect to the overall contribution of each individual modifier as well as the share associated with each construction. Yet, each two constructions per type of modifier are more likely to "cluster" regarding their

frequency, and these are not necessarily the two ARG-Cxx. This observation is the most pronounced with respect to QUANTITY modifiers where the ARG-Cxx are very similar in frequency with a lot of tokens, while the [*es*+*regnen*] construction has a low incidence of QUANTITY modifiers. On the other hand, it is the [*es*+*regnen*] and the [*Regen*+*geben*] construction, which converge frequencywise with regard to modifier STANCE. This trend is evident as well in the data for the overall most frequent TEMPORAL modifier, which is also the most frequent modifier for two of the constructions. More than half of the data overall (58%) encompasses at least one TEMPORAL modifier. These are, furthermore, fairly equally distributed proportionally between the three constructions in general: the two ARG-Cxx, [*Regen*+*fallen*] and [*Regen*+*geben*] each make up 36% of the data while the [*es*+*regnen*] construction lags by eight percentage points (28%). Modifiers of QUANTITY are the second most frequent and appear in more than one-third of the data; this is interesting, as QUANTITY is considered a peripheral element of the PRECIPITATION-frame. A closer look at the distribution of frequencies across the three constructions reveals that the overwhelming majority of QUANTITY modification occurs in the two ARG-Cxx. Here, the modifiers are furthermore relatively evenly distributed (47% [*Regen*+*geben*] vs. 44% [*Regen*+*fallen*]). In contrast, the [*es*+*regnen*] construction is strikingly infrequent and makes up the remainder of 9%. Modifiers of STANCE and QUALITY are the least frequent overall with 30% (STANCE) and 23% (QUALITY) respectively. Both QUANTITY and QUALITY, or rather [Quantity] and [Manner] are considered non-core elements of the PRECIPITATION.[81] As pointed out earlier, a case can be made to combine QUANTITY and QUALITY, as they certainly share similarities, but making use of this option, too, reveals differences between the three constructions: the two ARG-Cxx share similar frequencies (137 vs. 128 tokens respectively or 40% vs. 37%) where QUANTITY and QUALITY are combined, but the *es*+PRED-Cx amounts to only 77 tokens in total (23%). That means that modifications that are considered non-core elements, or peripheral in the PRECIPITATION-frame occur just as frequently, if not more frequently, than TEMPORAL modifications within the returns for the ARG-Cxx with the *Regen* 'rain' precipitate. For the *es*+PRED-Cxx, QUALITY and

[81] The Berkely Frame Net defines the frames [Manner] and [Quantity] as follows "Semantic Type: Manner: The way in which the precipitation falls. It DRIZZLED softly. Quantity: The Quantity is the amount of Precipitation that has fallen in a precipitation event or events. It is usually measured as the height that the Precipitation attains if prevented from further movement." (Source: framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Precipitation (last accessed: December 6, 2012)

QUANTITY are not as frequent as TEMPORAL modifiers but still amount to roughly two-thirds of the latter (77 vs. 121 tokens).

Lastly, STANCE is not considered in the PRECIPITATION-frame at all. Obviously, STANCE is not inherent in the weather event itself, but it is habitually supplied by the speaker. However, expressions of MANNER, which are peripheral elements of the PRECIPITATION-frame, are not neutral *per se*, but may carry semantic prosody (e.g., *It rained violently*, *destructive rainfalls*, *rejuvenating rains*) nonetheless. With respect to stancetaking the [*es+regnen*] construction (45%) and the [*Regen+geben*] construction (40%) "cluster" regarding frequency while *Regen+fallen* is the odd one out as it accounts for the remaining 15% of the data. Since both, the [*es+regnen*] construction and the [*Regen+geben*] construction, are associated more frequently with more informal discourse in the shape of the WEATHER TALK genre, it makes sense for them to more readily express STANCE than the [*Regen+fallen*] construction, which contributes the highest amounts of tokens to genres WEATHER REPORTS and METEOROLOGY, which appear to be more objective in their expressions of the weather situation, especially with respect to AVW and HWW. As a concluding remark, it should be noted that SPATIAL modifiers, which are considered core elements of the PRECIPITATION-frame, are by far not as frequent as modifiers of TIME or QUANTITY —especially with respect to the [*Regen+geben*] construction— although we might have expected them to return similar frequencies as modifiers of TIME. Instead, modifiers of PLACE occur predominantly within the *Regen+fallen* returns (43%), while the [*es+regnen*] construction roughly accounts for one-third of the LOCATIVES (32.5%).

3.1.5 Summary of Regen *'rain' constructions*

The preceding quantitative and qualitative analysis of MEs employing the 600 returned concordance lines for the three constructions *Regen+fallen*, *Regen+geben*, and the *es+regnen* has provided a comprehensive account of the constructional preferences of each of the three constructions and their interaction with a number of extra-propositional variables. While there is a great deal of overlap between the three constructions, each construction was also found to attract different variables to varying degrees. Generally, two constructions tend to be close together frequencywise while the third is the odd one out, representing with a notably higher or lower frequency than the other two constructions, depending on the variable and its levels.

First of all, a Pearson's chi-square test carried out for Table 34 (GENRE variable), Table 35 (TAM variable), Table 37 (SPATIAL) and Table 38 (NON- SPATIAL) revealed statistically highly significant results for each table. The chi-squared statistic (X-squared), the degrees of freedom (df), and the significance level (p-value) are listed in the following:

Table 34: X-squared = 89.278, df = 8, p-value = 6.515e-16

Table 35: X-squared = 117.69, df = 10, p-value < 2.2e-16

Table 37: X-squared = 184.39, df = 4, p-value < 2.2e-16.

Table 38: X-squared = 100.02, df = 6, p-value < 2.2e-16

For Table 37, the minor, or more peripheral spatial modifiers (goals, sources, and paths) were combined since otherwise the frequencies are too low for a reliable application of the chi-squared test for a significant distribution. Table 36, which shows the frequency and distribution of past tense differences across the three constructions and the precipitate *Regen* 'rain', does not allow for a reliable chi-square calculation, too, due to the low frequencies. However, a Fisher's exact test (an exact procedure especially for small-sized samples) returns a p-value of 0.2322, i.e., the distribution is not statistically significant.[82] The more general observations are summarized in the following:

The [*Regen+fallen*] construction seems to be mostly associated with genre type WEATHER TALK (46.5%), but to a notably lower degree than the [*es+regnen*] construction: not even half of the data are coded as WEATHER TALK. In comparison, the genre type METEOROLOGICAL represents more than one-fourth of the data. These concordance lines exclusively deal with reporting on the average weather of a region/country or explicate "how weather works". The remaining quarter is split between genres LITERARY and WEATHER REPORT. With respect to TAM factors, the majority of events are EXISTENTIAL or GENERIC—correspondingly, many of the TEMPORAL modifiers express larger time parcels, such as months, seasons, and annuality—followed closely by PAST tense. There is not much aspectual variation within the PAST tokens; instances of progressive aspect are few. Furthermore, there are no occurrences of FUTURE tense or IRREALIS. Regarding SPATIAL modification, the greater number of concordance lines encompass some type of LOCATION

[82] I would like to thank my examiner Antti Arppe for carrying out the Pearson's chi-squared test as well as the Fisher's exact test for Table 36 and sharing the results with me.

and a smaller number represents without any SPATIAL modification whatsoever. One-fifth of the returns also contain other SPATIAL mentions, particularly GOALS. This is a sizeable amount, especially in comparison with the other two constructions. As most of the figurative instances are analogies, the SPATIAL expressions in these returns do essentially not differ in nature from the literary instances. As with the *es*+PRED-Cx, the most frequent modifier is TEMPORAL, however, in contrast to the *es*+PRED-Cx, the second most frequent modifier marks QUANTITY. While the QUANTITY modifiers account for a large portion of modifiers, STANCE, again in contrast to the [*es*+*regnen*] construction, only plays a relatively minor role in the [*Regen*+*fallen*] data.

The query for the [*Regen*+*geben*] construction did not return a single concordance line being used figuratively but there are many false returns. This construction, too, is mostly attracted to the genre WEATHER TALK (65%), but it occurs also frequently in the LITERARY genre, followed by WEATHER REPORTS. The genre type METEOROLOGY only attracts the [*Regen*+*geben*] construction infrequently. With respect to the TAM variable, the majority of events are in PAST tense with pres_FUT and EXIS_GEN following each with about half as many instances as PAST tense. As opposed to the other two constructions, [*Regen*+*geben*] contains a large number of FUTURE tense inflections and generally appears to be the designated construction to lexicalize imminent or potential rain events. The bulk of the data, i.e., almost three-fourths, does not contain any SPATIAL modification. Where there is a mention of something SPATIAL, it tends to be a LOCATION, which is most frequently encoded by a deictic expression. Modifiers of TIME are again the most frequent modifiers, if only by a thin margin: TEMPORAL modifiers occur in half of the data followed closely by QUANTITY modifiers. The latter are regularly combined with expressions of STANCE. Moreover, a large number of QUANTITY modifiers are negators, i.e., they specify the absence of rain, i.e., ZERO rain. Expressions of QUALITY only play a minor role; the [*Regen*+*geben*] construction is the least likely of the three constructions to attract this type of modifier.

The [*es*+*regnen*] construction, too, seems to be attracted chiefly to the genre WEATHER TALK (80%), which can be more formal or rather informal, while *es*+*regnen*-type constructions rarely occur as WEATHER REPORTS. The majority of events are talked about in the PAST tense, with PRESENT tense running (a close) second. Overall, the PRED-Cx displays a large variety of aspectual use within the tenses with frequent use of progressive aspect and quasi-auxiliaries, which focus the attention predominantly on the incipient phase of the ME. Accordingly, many modifiers of time are rather idiosyncratic in that they specify unique points in time usually coinciding with other activities. Events that are expected in the future or are imminent, generally do not appear

to be encoded with the *es*+PRED-Cx. The same is true for "true" irrealis. This finding is likely to go together with the underrepresentation of WEATHER REPORTS, which obviously forecast events. More than half of the corpus returns do not contain any SPATIAL modification. Where there is a SPATIAL mention, it is usually a LOCATION. There are a handful of GOALS and SOURCES but not a single PATH. The most frequent modifier is TEMPORAL, followed by STANCE. The latter is not a component of the semantic FRAME for a precipitation event according to FRAME SEMANTICS but is frequently expressed in the *es*+PRED-Cx. QUANTITY, which is an optional component of the PRECIPITATION FRAME, is the least frequent NON-SPATIAL modifier to be expressed in utterances which employ the *es*+PRED-Cx and, furthermore, appears to be genre-specific. We shall see in the following sections whether the marking and modification found in this section for rain constructions hold true for the three constructions in combination with precipitate *Schnee* 'snow', as well as *Schneeregen* 'sleet' and *Hagel* 'hail'.

3.2 Constructions with the precipitate *Schnee* 'snow'

This section discusses the *Schnee* 'snow' precipitate, which is probably the second most frequent precipitate to occur naturally after *Regen* 'rain'. With respect to some constructions and their interactions with certain variables the *Schnee* 'snow' precipitate data sets appear to follow the tendencies of the *Regen* 'rain' precipitate. For example, the reporting on MEs predominantly in the PAST tense and within the genre type WEATHER TALK are similar, and even more pronounced regarding the latter. Having said that, between the constructions there are clear usage differences; for example, the WEATHER TALK instances in the [*Schnee*+*geben*] returns hail from very diverse range of sources. Within the *Schnee* 'snow' data sets themselves, the usage is much more literal than that of *Regen* 'rain', but, again, we shall see that each construction shares some of the qualities as their counterpart in the *Regen* 'rain' precipitate data sets: the [*Schnee*+*fallen*] construction encompasses the greatest variety of SPATIAL modifiers, the [*Schnee*+*geben*] construction has a propensity to lexicalize imminent or potential rain events, and the [*es*+*schneien*] construction mostly inflects for PRESENT or PAST tense with a penchant for aspect marking. The more subtle, but possibly also greater differences, may be found in the semantic make-up of the individual variables, which are elaborated on in the respective sub-sections.

3.2.1 The argument construction [Schnee+fallen] *'snow+'fall'*

This section surveys the ARG-Cx [*Schnee+fallen*] 'snow+fall'. The queries in the *deTenTen13* were carried out using the expressions in (ix) and (x) to account for the different word orders. The search for the below queries returned 17,771 concordance lines in total. The results are not as disparately distributed between word orders as those of the *es*+PRED-Cx with PREC1 being more frequent with more than 2,200 concordance lines more than PREC2.

(ix)	[word="Schnee"] [ ]{0,8} [lemma="fallen"] within <s/>	(N=9,989)
(x)	[lemma="fallen"] [ ]{0,8} [word="Schnee"] within <s/>	(N=7,782)

Forty-nine concordance lines from the PREC1 word order were removed as false returns but only 28 concordance lines from the PREC2 order. If that trend continues through the rest of the returns, then (ix) and (x) can be assumed to be even closer together in terms of frequency of results. Take note that the queries for this construction do not contain the second light argument *es* 'it', although *es fällt Schnee* 'it falls snow' is perfectly acceptable and used frequently. Having said that, this construction is only possible in the PREC2 configuration, but not PREC1 (**Schnee fällt es* 'rain falls it'). The frequencies for queries containing *es* 'it' are given in (xi) and (xii), from which we can see that the regular expression including the pronoun only returns 359 concordance lines, while the same search without any modifying material whatsoever, as in (xii), only returned 84 concordance lines in present tense, 31 concordance lines in preterite, and each seven concordance lines in *Perfekt* and past perfekt.

(xi)	[word="es"][lemma="fallen"][]{0,8}[word="Schnee"] within <s/> (N=359)
(xii)	simple concordance searches: *es fällt Schnee* ('it falls snow', N=84), *es fiel Schnee* ('it fell snow', N=31), *es ist Schnee gefallen* ('it has fallen snow', N=7), *es war Schnee gefallen* ('it had fallen snow', N=7)

The next subsections discuss the distribution, frequencies, and relative frequencies of the various variables itemized in Chapter 2.

3.2.1.1 The USAGE (LITERAL vs FIGURATIVE) and GENRE variables

Table 39 below gives the frequencies and distribution of the literal and figurative use for the ARG-*fallen*-Cx and precipitate *Schnee* ('snow').

Table 39. *Frequencies of literal and figurative USAGE for the ARG+*fallen*-Cx and precipitate* Schnee *'snow'*

USAGE / Word order	**PREC1** *Schnee+fallen*	**PREC2** *fallen+Schnee*	**Total**
LITERAL	100	100	200
FIGURATIVE	0	0	0
Total	100	100	200

The [*Schnee+fallen*] construction contains no instances of figurative use. There are two utterances which deal with artificial snow, which is "made fall" or 'let fall' (*fallen lassen*) by an intermediary, as in (94), which is taken from a report on a school performance.

(94) *Am Himmel schwebten **Rentiere** und **ließen <u>Schnee herunterfallen</u>**, so dass alle Tannenbäume weiß wurden*. (jag-bad-berleburg.de **'**Johannes Althusius Secondary School Bad Berleburg') '**Reindeer** were floating in the sky and **let <u>snow fall down</u>** so that all Christmas trees became white.'

These instances are obviously not genuine weather events but are designed to simulate MEs. Thus, they are not coded as FIGURATIVE. A total of eight concordance lines contain negation. The distribution and frequencies of GENRE for the [*Schnee+fallen*] constructions are given in Table 40. The last column shows the relative frequency of the total per 200 concordance lines. With respect to the aggregate frequencies of variable GENRE, WEATHER TALK is the principal genre with more than half of the corpus returns (N=112; 56%) followed by genre types METEOROLOGY and WEATHER REPORT, which are exactly evenly distributed (N=32; 16%). The LITERARY genre has the lowest incidence (N=22 or 11%). Lastly, the category OTHER amounts to two concordance lines (1%). These are instances where human agents "let snow fall", see ex. (94) above.

Table 40. *Raw and relative frequencies for the* GENRE *variable across the ARG+*fallen*-Cx and precipitate* Schnee *'snow'*

Word order/ GENRE	PREC1 *Schnee+fallen*	PREC2 *fallen+Schnee*	Total	Relative Frequency /200
WEATHER TALK	48	64	**112**	**56%**
WEATHER REPORT	16	16	**32**	**16%**
METEOROLOGY	15	17	**32**	**16%**
LITERARY	19	3	**22**	**11%**
OTHER	2	0	**2**	**1%**
FIGURATIVE	0	0	**0**	**0%**
Total	**100**	**100**	**200**	**100%**

The frequencies are distributed fairly evenly between the two word orders. The one result that stands out is the relatively high number of LITERARY tokens in the PREC1 order. In return, PREC2 has a comparatively higher number of WEATHER TALK tokens. The METEOROLOGICAL genre is mainly concerned with the average weather in certain regions (N=24) rather than how weather works (N=6). The concordance lines from the LITERARY genre mainly come from fiction (N=20) with only two instances from religious texts. The heterogenous distribution between the two word orders is interesting, since the majority of LITERARY tokens occur in the PREC1 order. This can mostly be ascribed to the utterances containing modals, future marker *werden* ('is going to'), the hypothetical marker *würde* ('would'), and the presence of quasi-auxiliaries (N=6), all of which move the non-finite verb *fallen* 'fall' to a sentence final position. The diversity of finite verbs in these cases is also reflected in the diversity of tenses they are inflected for. The TAM variable is discussed in the next section.

3.2.1.2 The TAM (TENSE, ASPECT, MOOD) variable

The distribution and frequencies of the set of TAM variables for the [*Schnee+fallen*] construction are given in Table 41. Examining the total count, there is a noticeably homogenous distribution between concordance lines in PAST tense, pres_FUT, and EXISTENTIAL/GENERIC. PAST tense is slightly more frequent than present_FUTURE and EXISTENTIAL/GENERIC (hereafter EXIS_GEN) tokens, which are evenly distributed. However, present_FUTURE, FUTURE and IRREALIS tokens subsumed under "future intent" clearly have the highest incidence of all TAM levels (N=72; 36%). The majority of the "future intent" returns contain a modal verb (N=38). PRESENT tense utterances are about half as frequent as PAST, present_FUTURE, and EXIS_GEN tokens and contain 17 instances of progressive aspect. Due to word order preferences, all FUTURE and present_FUTURE tokens with a modal verb belong to the PREC1 word order.

Table 41. *Raw and relative frequencies of the* TAM *variable across the ARG+*fallen*-Cx and precipitate* Schnee *'snow'*

Word order/ TAM	**PREC1** *Schnee+fallen*	**PREC2** *fallen+Schnee*	**Total**	**Relative frequency** /200
PAST	7	47	**54**	**27%**
PRESENT /PROG	6/9= 15	1/8= 9	**24**	**12%**
pres_FUT	36	14	**50**	**25%**
FUTURE	16	0	**16**	**8%**
IRREALIS	6	0	**6**	**3%**
EXIS_GEN	20	30	**50**	**25%**
Total	100	100	**200**	**100%**

Within the PAST returns, there is not a single instance of progressive aspect, *Perfekt*, or past perfect; all PAST instances are in the preterite. Nonetheless, a simple search for "*Schnee gefallen*" ('snow+past participle of fall') in the *deTenTen13* returns 2,707 results, as in (95). This is interesting, considering that the query in (ix) returned 9,989 concordance lines in total and that, apparently, none of these contain any instance of *Perfekt* in the downloaded sample.

(95) *Wenn auch recht spät, so* ***ist nun doch Schnee gefallen*** *– der Winterurlaub 2007 scheint gerettet.*
(versicherung-weblog.de **'**insurance-weblog')
'Snow has fallen, albeit quite late – the 2007 winter holiday seems saved.'

Based on word order, it is striking that 47 out of 54 PAST instances feature the PREC2 word order. Every single of these concordance lines contains a locative or temporal modifier at the beginning of the sentence. In contrast, four of the PREC1 order utterances contain a finite quasi-auxiliary, which moves verb *fallen* 'fall' to a sentence final position, and three instances contain verb *lassen* 'let', as in example (94). Regarding quasi-auxiliaries *start*/*begin* and *stop*, there are six instances in total (*beginnen* 'begin', N=4; *anfangen* 'start', N=1; and *nicht aufhören* 'not stop' N=1), which all occur in concordance lines coded as LITERARY and furthermore come from fiction.

3.2.1.3 The SPATIAL modifier variable

The distribution and frequencies of the SPATIAL modifier variable for the [ARG+*fallen*] construction are given in Table 42. The last column shows the relative frequencies of the total per 200 concordance lines. The complex SPATIAL mentions were split up and counted separately. There are five combinations of SPATIAL modifiers overall, which are listed right below the table. [*Schnee+fallen*] overwhelmingly contains LOCATIONS (N=122; 61%), of which the majority (N=101; 50.5%) directly modify the ME. Only 21 LOCATIONS indirectly modify the ME in the wider context.

Less than one-third of the data (N= 59; 29.5%) do not contain any SPATIAL modifiers whatsoever (ZERO). In a similar manner to the [*Regen+fallen*] constructions, other SPATIAL mentions, like GOAL (G), SOURCE (S), and PATH are common, albeit not as frequent as within the [*Regen+fallen*] returns. Here, too, GOALS are the most frequent (N=13), followed by SOURCES (N=10), and PATHS (N=3). Five of these minor SPATIAL modifiers occur as complex paths in combination with a LOCATION (L). The nature of the SPATIAL modifiers is elaborated on below.

Table 42. *Raw and relative frequencies of the SPATIAL modifier variable across ARG+*fallen *and the precipitate* Schnee *'snow'*

Word order/ SPATIAL modifier	**PREC1** *Schnee+fallen*	**PREC2** *fallen+Schnee*	**Total**	**Relative frequency** /200
LOCATION	47	75	**122**	**61%**
GOAL	8	5	**13**	**6.5%**
SOURCE	7	2	**9**	**4.5%**
PATH	3	0	**3**	**1.5%**
ZERO	38	21	**59**	**29.5%**
Total	**103**	**103**	**206**	**103%**
COMBINATIONS*	2	3	5	2.5%

*PREC1= each one L+G, L+S+G; PREC2= L+G (N=2), L+S (N=1)

The SPATIAL modifiers can be grouped into eight categories: "location", "in the mountains", "altitude", "deictics", "cardinal direction", "source/goal", "combos", and "other". The most frequent category is "location" with at least 28 tokens; these can be as simple as *auf Sylt* 'on [the island of] Sylt' and as complex as *auf einigen hochgelegenen Passstraßen im Pindusgebirge* 'on some high pass roads in the Pindus Mountains'. The second most frequent category are "sources and goals" (N=22), as can be gathered from Table 42. The most frequent type of source is *vom Himmel* 'from the sky' (N=7). Goals, as we have seen with other constructions, are more varied and can operate on a more macro level (*auf die Erde* 'onto the ground/the Earth') or more focused on specific objects (*auf das Fenster* 'onto the window', *auf einen kahlen Baum* 'onto a bare tree'). The three instances of paths could be added to sources and goals, they constitute instances of von *oben nach unten* lit. 'from up to below/down', *herunter* 'downwards', and *durch das matte Grau* 'through the matte grey". "Deictics" have a comparatively high incidence as well (N=19), consisting mostly of *hier* 'here', *dort* 'there', and *draußen* 'outside'. Another type of special LOCATIVES refer to the mountains (N=13), e.g., *in den Bergen* 'in the mountains' or *wo es gebirgig ist* 'where it is mountainous', or which identify vertical height or "altitude" (N=10), e.g., *oberhalb 300 bis 800 Meter* 'above 300–800 meters' or

in höheren Lagen 'at higher altitudes'. The last two categories could be grouped together, the main difference is that "altitude" tokens contain an explicit numeral. It is not surprising that many of the SPATIAL modifiers refer to mountainous regions, altitude, and even specific places located in the mountains, or which are known for skiing, like the Mutnovsky volcano in the Russian far east (e.g., *im Mutnovskigebiet* 'in the Mutnovsky ski region') since higher elevations tend to receive more snow. Slightly more frequent than "altitude" are modifiers of "cardinal direction" (N=12), such as *im Westen und Süden* 'in the west and the south'. Lastly, there a few combinations such as *bei uns in NRW* 'here in NRW', which combines a deictic adverb with a location, or *nur in den Bergen und Richtung Norden* 'only in the mountains and the northern direction', which combines the mountains with a cardinal direction. Implied locations are such as *wenn wir die Marchfelder Verwandten besuchten* 'when we were visiting relatives in the Marchfeld region'. This utterance also contains an element of temporality, which belongs to the NON-SPATIAL modifiers discussed in the next section.

3.2.1.4 The NON-SPATIAL modifier variable (TEMPORAL, QUANTITY, QUALITY, STANCE)

The distribution and frequencies of various MODIFIERS for the ARG+*fallen*-CX are given in Table 43. The last column shows the relative frequency of the total per 200 concordance lines., there are 273 NON-SPATIAL modifiers and almost three-fourth of the returns (N=146; 73%) contain at least two different types of modifier, NON-SPATIAL or SPATIAL. Only three concordance lines (1.5%) do not contain any type of modifier (ZERO).

Table 43. *Raw and relative frequencies of the NON-SPATIAL modifier variable across the ARG+*fallen*-Cxx and the precipitate* Schnee *'snow'*

Word order/ NON-SPATIAL modifier	**PREC1** *Schnee+fallen*	**PREC2** *fallen+Schnee*	**Total**	**Relative frequency** /200%
TEMPORAL	60	70	**130**	**65%**
QUANTITY	32	38	**70**	**35%**
QUALITY	10	28	**38**	**19%**
STANCE	22	13	**35**	**17.5%**
Total	**124**	**149**	**273**	**136.5%**
ZERO	3	0	3	1.5%
COMBINATIONS	65	81	146	73%
SPATIAL	54	65	119	59.5%

The most frequent modifier is TEMPORAL with 130 concordance lines (65%) containing at least one expression of time (N=70; 35%). Modifiers of QUALITY and STANCE are the least frequent but

are fairly evenly distributed. The following subsections elaborate on each NON-SPATIAL MODIFIER individually.

Modifiers of TIME. Almost two-thirds of the [*Schnee+fallen*] returns contain at least one expressions of time, which makes them not just the most frequent type of NON-SPATIAL modifier, but they are also highly varied regarding their semantics. This is true not just with respect to types but also larger categories that the types can be grouped into. These categories are "deictics", "time of day", "day of the week", "month", "season", "exact date/day" (e.g., *am 1. Juni 2005* 'on June 1, 2005', *am Heiligen Abend* 'on Christmas Eve'), "time period/duration", "idiosyncratic", and "other/combinations", where the last few are the least frequent, except for the catch-all category "other/combinations". The latter encompasses instances where more than one temporal expression is given within the same utterance, such as *Samstag Mittag … innerhalb kürzester Zeit* 'Saturday midday … within the shortest amount of time'. The most frequent expression is *im Winter* 'in (the) winter' (N=13), including slight variations, such as *über den Winter* 'throughout winter'. Again, this finding is not surprising since wintertime is the default season for snowfall. Correspondingly, the months mentioned range from *Oktober* to *April*. For the "time of day" category, nighttime seems to be a common modifier, presumably due to the fact that the result of an overnight snowfall, i.e., snow on the ground, is usually more evident than rainfall. As we have already seen with *Regen* 'rain', "deictics" tends to be a category with repetitive tokens. Here, "deictics" contains several tokens of *heute* 'today' (N=5), *bald* 'soon' (N=4), *gestern* 'yesterday' (N=3), as well as unique items such as *nun* 'at the moment', *jetzt* 'now', and *am nächsten Tag* 'the next day'.

Modifiers of QUANTITY. QUANTITY modifiers are almost as frequent as modifiers of QUALITY and STANCE together and can be grouped into five semantic categories. First of all, there is quantity "zero", i.e., "negations" with *kein* 'no'. Furthermore, there are "small quantities", such as *etwas* 'a little/some' (N=8), *wenig* 'a little/some' (N=6), *kaum* 'hardly any' and countable numbers of snowflakes, such as *drei/ein paar Flocken* 'three/a couple of flakes' (N=3) and there are "large quantities", i.e., variations of *viel* 'a lot/much' (N=10), such as *sehr viel* 'very much', *richtig viel* 'really much', *derart viel* 'so much', *zu viel* 'too much', *recht viel* 'quite a lot', as in (96).

(96) ***so viel …, dass*** *Autobahnen und wichtige Verkehrsanbindungen vollständig blockiert waren*
'**so much** …, **that** motorways and important transport connections were completely blocked'

The fourth category is “quantity in numbers”, i.e., the quantity is identified as a metric measurement, such as in centimeters or meters. The height is not limited to either low or high numbers, but the 10 tokens range in declaring snowfall between *5cm* and *313cm*, as illustrated in (97). Additionally, approximate numbers tend to be expressed those phrases such as *mehrere Meter* ‘several meters’ or *halber Meter* ‘half a meter’.

(97) *Als massiv würde ich einen Wintereinbruch dann bezeichnen,* ***wenn innerhalb von ein oder zwei Stunden 30 cm Schnee fallen****!*
‘I would call a sudden start of winter massive **when 30 cm of snow fall within an hour or two**!’

Lastly, the catch-all “other” encompasses instances where it is not quite clear whether the modifier means *a lot* or *a little*.

Modifiers of QUALITY. Modifiers of QUALITY are the second least frequent type of NON-SPATIAL modifier (N=38) and can be separated into two larger semantic categories. The first category describes the “kind of snow“ (N=18), usually *der erste* ‘the first’, which is obviously specific to snow. Speakers do not tend to proclaim the first rain, sleet, or hail of the season, year, or winter. Secondly, there are 12 unique tokens, which refer to characteristics of the snow, such as *nass* ‘wet’, *dicht* ‘dense’, *leicht* ‘light’, *neu* ‘new’, *weiß* ‘white’, as well as the manner of the snow falling, such as the path it takes: *von rechts nach links statt von oben nach unten* lit. ‘from right to left instead of from above to down’ and also *in verschiedenen Intensitäten* ‘in different intensities’. Lastly, there are 21 concordance lines which contain *als Schnee fallen* ‘fall as snow’. However, this is the quality of the *precipitation* rather than the snow itself. More than half of these utterances specify snow as the only precipitate falling, but no other precipitates, as in (98) below.

(98) ***Da der Niederschlag überwiegend als Schnee fallen sollte****, können bis zum Abend 3-15 cm Neuschnee erwartet werden, örtlich gar bis 25 cm*. (wetterprognose-wettervorhersage.de ‘weatherprognosis-weatherforecast‘)
‘**Since the precipitation is expected to fall mainly as snow**, 3-15 cm of fresh snow can be expected by the evening, locally even up to 25 cm.’

Other examples mention the precipitate rain or rain falling “as snow” above a certain altitude. The majority of *als Schnee fallen* ‘fall as snow’ utterances belong to the genre types WEATHER REPORT (N=10) and METEOROLOGY (N=8); only three tokens are coded as WEATHER TALK.

Modifiers of STANCE. Expressions of stancetaking follow the standard practice of expressing the categories “positive”, “negative”, “mirative”, and “other” where all four categories are close together with respect to frequency. STANCE is the least frequent modifier

(N= 35) and mostly conveys a negative attitude (N=11), as in (99). Expressions of surprise, such as *sogar* 'even' but also by literally using variations of *Überraschung/überrascht* 'surprise/surprised' have the second highest incidence (N=9).

(99) *In den nächsten Tagen* ***soll weiter Schnee fallen****.* ***Das dürfte*** *bei den Autofahrern* ***nicht gerade für Freude sorgen.*** (kfzticker.de 'vehicleticker')
'**More snow is supposed to fall** in the next few days. **That shouldn't exactly make** drivers **happy**.'

A positive stance is expressed in eight concordance lines, which contain no negation, i.e., the speaker is genuinely enjoying the snow or hoping for it. Catch-all category "other" is the least frequent with seven tokens. The next section details the distribution and frequencies of variables across the second ARG-construction, *Schnee+geben* 'snow+give'.

3.2.2 The argument construction [*Schnee+geben*] *'snow'+'give'*

This section discusses the ARG-Cx *Schnee+geben* 'snow'+'give'. The queries in the *deTenTen13* were carried out using the expressions in (xiii) and (xiv) to account for the different word orders. The search for the below queries returned 3,600 concordance lines in total. Take note that these queries contain pronoun *es* 'it', so that there are two arguments in this construction, one which carries all the semantic weight (*Schnee* 'snow') and one which is "empty" (*es* 'it'). As opposed to the [*Schnee+fallen*] construction, the pronoun es 'it' is obligatory in this construction and may not be left out. For convenience, the construction is still referred to as the [*Schnee+geben*] construction.

(xiii) [word="es"][]{0,4}[lemma="geben"][]{0,8}[lemma="Schnee"] within <s/> (N=1,257)

(xiv) [word="es"][]{0,4}[lemma="Schnee"][]{0,8}[lemma="geben"] within <s/> (N=2,343)

Still, queries without pronoun *es* 'it' were carried out, as in (xv) and (xvi), simply to gauge possible differences in frequency, since there is no restricted number of items between the word and the first lemma, as in (xiii) and (xiv). The queries below returned 17,145 concordance lines in total as well as the overall most evenly distributed frequencies between the two word orders, compared to the [*Schnee+fallen*] construction and particularly the [*es+schneien*] construction but also compared to the queries in (xiii) and (xiv).

(xv) [word="Schnee"][]{0,8}[lemma="geben"] within <s/> (N=8,976)
(xvi) [lemma="geben"][]{0,8}[word="Schnee"] within <s/> (N=8,169)

There is quite some overlap between (xv) and (xv), since the position of the pronoun is not restricted. Accordingly, *es* 'it' can appear before (*es gibt Schnee* 'it gives snow'), after (*Schnee gibt es* 'snow gives it'), or in the middle of the verb-noun complex (*gibt es Schnee* 'gives it snow'). From scanning the first few pages of the search results for the PREC2 order, the pronoun seems to appear most frequently between the verb and the noun, whereas for the PREC1 order it tends to occur either before or after the noun-verb complex. Either way, since the [*es*+*geben*] constructions are ubiquitous, the queries also returned a fair share of false returns. Only ten false returns had to be removed from the in (xiii) search results, however, from the (xv) search results, 65 concordance lines had to be removed as false returns. In most of these, verb *geben* 'give' relates to a different object than *Schnee* 'snow'.

3.2.2.1 The USAGE (LITERAL VS FIGURATIVE) and GENRE variables

Table 44 gives the frequencies and distribution of the literal and figurative uses for the [ARG+*geben*] construction with precipitate *Schnee* 'snow'. Strikingly, and just like within the [*Schnee*+*fallen*] returns, there are no figurative uses within the [*Schnee*+*geben*] data subset. Furthermore, there are no instances of "other" where there is no actualized ME taking place, such as intermediaries letting artificial snow fall, as in (108) above, or similar instances.

Table 44. *Frequencies of literal and figurative USE for the ARG+*geben*-Cx with the precipitate* Schnee *'snow'*

USAGE/ Word order	**PREC1** *Schnee+geben*	**PREC2** *geben+Schnee*	**Total**
Literal	100	100	**200**
Figurative	0	0	**0**
Total	**100**	**100**	**200**

There are overall 38 instances of negation, such that almost 20% of the returns contain some sort of negative marking, mostly with the negative article *kein* 'no'. There are also two instances of *weder … noch* 'neither … nor'. Out of all the negated utterances, 15 carry negative semantic prosody where the speaker laments the absence of snow. Six concordance lines carry positive semantic prosody where the speaker is glad that it did not snow or they wish for it not to snow. The distribution and frequencies of the GENRE types for the [*Schnee*+*geben*] construction are given in Table 45. The last column shows the relative frequency of the total per 200 literally used

concordance lines. With respect to the aggregate frequencies of the GENRE variable, WEATHER TALK is the principal genre by a wide margin with more than four-fifth of the returns (N= 163; 81.5%). There are a few instances of returns from the METEOROLOGY genre (N=18; 9%), closely followed by concordance lines from the LITERARY genre (N=13; 6.5%). WEATHER REPORTS have the lowest incidence with only six tokens (3%). There are no utterances that do not belong to any of these four categories.

***Table 45**. Raw and relative frequencies of the GENRE variable across the ARG+*geben*-Cx with the precipitate* Schnee *'snow'*

Word order / GENRE	**PREC1** *Schnee+geben*	**PREC2** *geben+Schnee*	**Total**	**Relative frequency** /200
WEATHER TALK	78	85	**163**	**81.5%**
WEATHER REPORT	3	3	**6**	**3%**
METEOROLOGY	11	7	**18**	**9%**
LITERARY	8	5	**13**	**6.5%**
Total	**100**	**100**	**200**	**100%**

Interestingly, with respect to frequency, the genre types resemble the [*es+schneien*] construction much more closely than the [*Schnee+fallen*] construction. Scanning the WEATHER TALK tokens, they come from a wide variety of source topics: reporting on community news (fire brigade, construction work), events (ice bathing, music), sports (skiing, motor racing, running), private blogs, "talking about the past" (childhood, Ottonion dynasty), travel, insurance, animals (dog owners), photography/film, fashion, gardening, etc. Corresponding with the diversity of topics is the diversity of TAM markers employed for the WEATHER TALK genre, which are surveyed in the next subsection.

3.2.2.2 The TAM (TENSE, ASPECT, MOOD) variable

The distributions and frequencies of the TAM variable for the [*Schnee+geben*] construction are given in Table 46. Examining the frequency aggregates for each TAM factor, the majority of concordance lines are in the PAST tense (N=86; 43%). The "future intent" factors (pres_FUTURE, FUTURE, and IRREALIS) together have the second highest incidence (N=56, 28%) despite the low incidence of WEATHER REPORTS, which have the tendency to co-occur with "future intent" inflections. Where "future intent" and weather talk interact the speaker usually cites the weather report and that 'it is supposed to give snow' (*es soll* [*ja*] *Schnee geben*), as in (106) in Section 3.2.2.4. Accordingly, there are 13 instances of *soll* 'supposed to' in addition to other, less

frequent types of modals. FUTURE tense alone, which generally occur with a low incidence, amounts to 14 tokens. These mostly interact with genre type WEATHER TALK, but they also occur in genre types METEOROLOGY and LITERARY. There is a fairly even distribution between concordance lines in PRESENT tense (N=32; 16%), and EXISTENTIAL/GENERIC use (N=26; 13%) with a noticeable absence of progressive aspect (N=1). The returns do not contain a single quasi-auxiliary either.

***Table 46**. Raw and relative frequencies of the TAM variable for the ARG+*geben*-Cx with the precipitate* Schnee *'snow'*

Word order / TENSE	**PREC1** *Schnee+geben*	**PREC2** *geben+Schnee*	**Total**	**Relative frequency** /200
PAST	33	53	**86**	**43%**
PRESENT	14	18	**32**	**16%**
pres_FUTURE	22	11	**33**	**16%**
FUTURE	12	2	**14**	**7%**
IRREALIS	4	5	**9**	**5%**
EXIS_GEN	15	11	**26**	**13%**
Total	100	100	**200**	**100%**

With respect to variation within PAST, the PREC2 order, which is much more frequent than the PREC1 order, only contains a single instance of *Perfekt*; supposedly, mostly due to word order preferences. The PREC1 order is much more diverse and contains 15 instances of *Perfekt* and one instance of past perfect.

3.2.2.3 SPATIAL modifier variable

The distributions and frequencies of the SPATIAL modifier variable for the [*Schnee+geben*] construction are given in Table 47. The last column shows the relative frequencies of the total per 200 concordance lines. The SPATIAL mentions mark either LOCATIONS or ZERO modification. LOCATIONS are slightly more frequent with more than half of the data (N=108; 54%) where 68 LOCATIONS directly modify the ME. In another 40 concordance lines, the LOCATION is given in the wider preceding or following context. The remainder (N=92; 46%) do not contain any SPATIAL modification whatsoever (ZERO). There are no GOALS, SOURCES, or PATHS, and accordingly there are no combinations of different SPATIAL modifiers. The types of LOCATIONS are further elaborated on below.

Table 47. *Raw and relative frequencies of the* SPATIAL *modifier variable for the ARG+*geben*-Cx with the precipitate* Schnee *'snow'*

Word order / SPATIAL modifier	**PREC1** *Schnee+geben*	**PREC2** *geben+Schnee*	**Total**	**Relative frequency** /200
LOCATION	57	51	**108**	**54%**
SOURCE	0	0	**0**	**0%**
GOAL	0	0	**0**	**0%**
PATH	0	0	**0**	**0%**
ZERO	43	49	**92**	**46%**
Total	**100**	**100**	**200**	**100%**

The modifiers of LOCATION can be grouped into five semantic categories with at least five tokens each. The most frequent category is "location" itself, which mention specific places and encompass *auf der Erde* 'on Earth', countries, cities, islands, regions, and more focused/zoomed in LOCATIONS, such as *auf der Piste* 'on the slopes'. Again, it is not always easy to tease apart whether the speaker is referring to snow on the ground or the actualized ME; locations such as the last mentioned in all probability refer to snow on the ground. The context usually decides the most likely scenario. The second larger group, although it has a much lower incidence than "locations", is "deictics" (N=19) and is exclusively made up of the types *hier* and *bei uns/bei mir* ('here' and 'at ours/at mine') as well as *dort* and *da* 'there', i.e., mostly adverbs. The two smaller groups, with roughly a handful of instances, are "cardinal directions", such as *im Osten des Landes* 'in the east of the country' and relative clauses introduced by relative pronoun *wo* 'where' (N=5), as in (102) below.

3.2.2.4 The NON-SPATIAL modifier variable (TEMPORAL, QUANTITY, QUALITY, STANCE)

The distributions and frequencies of various NON-SPATIAL modifiers for the [*Schnee+geben*] construction are given in Table 48. The last column shows the relative frequency of the total per 200 concordance lines. Overall, there are 294 NON-SPATIAL modifiers and roughly two-thirds of the returns contain at least two different types of modifier, NON-SPATIAL or SPATIAL (N=128; 64%). Eleven concordance lines (5.5%) do not contain any modifier whatsoever (ZERO).

Table 48. *Raw and relative frequencies of the* NON-SPATIAL *modifier variable within the ARG+*geben*-Cx and the precipiate* Schnee *'snow'*

Word order / NON-SPATIAL	**PREC1** *Schnee+geben*	**PREC2** *geben+Schnee*	**Total**	**Relative frequency** /200
TEMPORAL	58	53	**111**	**55.5%**
QUANTITY	50	49	**99**	**49.5%**
QUALITY	3	7	**10**	**5%**
STANCE	42	32	**74**	**37%**
Total	**153**	**141**	**294**	**147%**
ZERO	6	5	11	5.5%
SPATIAL modifier	37	33	70	35%
COMBINATIONS	70	58	128	64%

All types of NON-SPATIAL modifier appear to be fairly frequent with a proportion of between 35% and 55.5% each, except for modifiers of QUALITY, which are even less frequent than ZERO modification (5.5%), if only marginally. More than half of the data (N=111; 55.5%) contain at least one modifier of TIME and almost half of the data contain a modifier of QUANTITY (N=99; 49.5%). Expressions of STANCE are present in more than one third of the returns (N=74; 37%). The most frequent stand-alone modifier is TEMPORAL (N=13). The most frequent combination of NON-SPATIAL modifiers is that of TEMPORAL+ STANCE (N=24), often complemented by a modifier of QUANTITY (TEMPORAL+ STANCE +QUANTITY: N=14). The following subsections deal with each type of NON-SPATIAL modifier individually.

Modifiers of TIME. TEMPORAL modifiers do not just occur with many tokens (N=111), but the types, too, exhibit ample variety. The seven semantic categories that have emerged are "deictics", "frequency", "months", "season", "day of the week", "exact date", and "other". The most frequent two categories are "deictics" and "season" with 29 and 24 tokens respectively, although there is some overlap with other categories for both. The most prolific type of season is clearly and unsurprisingly *Winter* 'winter', as in (100), which informs the reader about the general climate in Ireland (genre type METEOROLOGY), but there are also the unique items of *Sommer* 'summer' and *Herbstanfang* 'beginning of fall'. A few of the *Winter* tokens also refer to past times with speakers noting that the winters in the past used to be richer in snow. These musing are furthermore accompanied by tokens such as *früher* 'past' (as in belonging to a former time), *damals* 'back in the days', and *wie in alten Zeiten* 'like in the old days'.

(100) ***Die Winter*** *sind mild und* ***es gibt in einzelnen, sehr seltenen Fällen Schnee****.* (farmarbeit.de 'farmwork')
'The winters are mild and **it gives in isolated, very rare cases, snow**.'

The items referring to "months" (N=11) generally involve the months between *Oktober* and *April*, but there is also one type mentioning *end of May* and *beginning of June*. Regarding "specific days/times" (N=7), these often refer to Christmas, the winter holidays, or days exceptional to the author, such as their birthday or wedding day. Example (101) refers to *Karfreitag* 'Good Friday' and is coded as genre WEATHER REPORT; despite the URL, it comes from a thread solely dedicated to reduplicating weather forecasts.

(101) *Am schlechtesten ist das Wetter* ***am Karfreitag, es gibt noch mal ordentlich neuen Schnee****, wieder im Norden und Osten, in der Mitte des Landes sogar bis zu zehn Zentimeter hoch*. (doctorsdiaryfanforum.de)
'The weather is the worst **on Good Friday, it gives a lot of new snow again**, again in the north and east, in the middle of the country even up to ten centimetres high.'

Although there are only 15 tokens in the "frequency" category, adverb *wieder* 'again' and its modifications (e.g., *irgendwann wieder* 'someday again') occurs with six tokens. There are eight concordance lines where an "exact date" is given (day/month/year) and sometimes additionally the day of the week. Lastly, the category "other" mostly contains unique items but also combinations such as (102) in the next section (*immer Schnee, so viel Schnee an einem Tag* 'always snow, so much snow in one day'), and time spans (*innerhalb von drei Tagen* 'within three days').

Modifiers of QUANTITY. Modifiers of QUANTITY have the second highest incidence and appear in nearly of half of the data. Semantically, the modifiers can be sorted into five bigger categories: "negation" (*keinen* 'no'), "small quantities" (*kaum* 'hardly'), "large quantities" (*viel* 'a lot'), "numbers" (*über einen halben Meter* 'more than half a meter'), and "other" (e.g., *genügend* 'enough'). The largest category is "negation" with 38 items. This finding, as well as the overall frequency of quantitative modifiers, makes the [*Schnee+geben*] results very similar to the [*Regen+geben*] results. The marking of large quantities contain the second highest incidence (N=35) of QUANTITY modifiers. More than half of these are variations of type *viel* 'a lot'. The category "small quantities" (N=17) is roughly half as frequent as "large quantities". These items are mostly variations of *wenig* 'little' and *ein bisschen* 'a little/some'. The remainder of tokens is split between "numerals" (*über einen halben Meter* 'more than half a meter') and "other"

(*genügend* 'enough' N=3, *mal wenig, mal viel* 'sometimes a little, sometimes a lot'), where the former refers to the quantity of snow on the ground after it has fallen. QUANTITY modifiers were easily categorizable, despite the fairly large number of tokens, as the types were recurrent and simple. The most verbose or complex modification of quantity is given in (102).

(102) *Und immer weiter voran, immer weiter – wohin?! In immer schrecklichere Öden,* ***wo es nur Schnee gab, immer Schnee, so viel Schnee an einem Tage, wie er in den ganzen zwanzig Jahren seines Lebens zusammen nicht gesehen hatte****.* (sophie-lange.de)
'And further and further, always further – where to?! Into ever more terrible wastelands, **where it gave only snow, always snow, more snow in one day than he had seen in all twenty years of his life**.'

Unsurprisingly, this example is from the LITERARY genre, more specifically from a narration named *Das Kriegsandenken* 'The War Token', written by Clara Viebig in 1905, in which the author vividly describes a trek of French prisoners of war (POW) struggling through the snow during the Franco-Prussian war. The fact that Viebig's POW-protagonist experiences *falling* snow, rather than snow on the ground, or possibly both, is spelled out a few paragraphs earlier in the narration where it says *und es schneit* 'and it is snowing.'. It is possible that the high number of QUANTITY modifiers, in a way, preempts modification of QUALITY, which have a very low incidence, discussed in the next subsection.

Modifiers of QUALITY. There are hardly any modifiers of QUALITY or MANNER in the [*Schnee+geben*] corpus returns. A mere 10 concordance lines (5%) contain a QUALITY modifier. These modifiers are *den ersten* 'the first', den *ersten leichten* 'the first light', den *ersten richtigen* 'the first real' as well as *richtigen* 'real', *echten* 'real/genuine', and *neuen* 'new' (all determiners and adjectives are masculine and in accusative case). In another example, the snow is described as *kalt und feucht* 'cold and wet'. Lastly, one modification of QUALITY is given as a relative clause:

(103) "*Dananä, Dananä,* ***es gibt an Schnee, der apert nimmermeh***!" (3sat.de 'public service television channel)
'Dananä, Dananä, **it gives a snow that melts nevermore**.'

The exclamation in the example above is a curse from an Alpine legend that dooms the mythological golden city of *Dananä*, which is believed to be located in Tyrol in Austria, to experience eternal winter. This ultimately leads to the demise of the city and its being lost in perpetual ice. The verb *apern* 'become free of snow/melt' is a regional expression used in

Austria and Switzerland (cf. the *DWDS*).[83] There are no instances that describe the manner of falling, such as *severely* or *incessantly*, i.e., a manner that can "double" as a modifier of QUANTITY. There also is only one instance of reduplication, given in (102), which differs from other types of reduplication in the data since each mention of snow in the example comes with its own modification (1. *nur Schnee* 'only snow', 2. *immer Schnee* 'always snow', 3. *so viel Schnee an einem Tage wie*… 'so much snow in one day than …') while in other instances of reduplication in the returned data, either the whole construction is reduplicated (*es schneit*, *es schneit* 'it's snowing, it's snowing') or only the predicate or argument (*es schneit*, *schneit*, *schneit* 'it snows, snows, snows'). However, for the *Schnee* 'snow' precipitate, these reduplications are reserved for the *es*+PRED-Cx. Example (102) is the only example in both ARG-Cxx where some kind of reduplication illustrates the seemingly inescapable mass of snow.

Modifiers of STANCE. More than one third of the returns for the [*Schnee+geben*] construction contain expressions of STANCE and more than one third of the tokens coded as STANCE (N=27) convey positive feelings about snow having fallen, as in (104). Another 13 concordance lines express disappointment that there was no snow or hardly any snow. Only one utterance communicates relief over snow having failed to materialize, as in (105). Fifteen utterances express negative feelings about snow having fallen. Lastly, there are 18 concordance lines, which indicate some "other" STANCE, such as mirativity, i.e., surprise or disbelief, but also evidentiality as in (106), where the speaker 'heard a rumour'.

(104) *Nur im Rheinland ist kaum Schnee... deswegen* ***freue mich so unbändig, wenn es mal so viel Schnee gibt.*** (sichtwiese.blogspot.de 'view meadow')[84]
'Only in the Rhineland is there hardly any snow… that's why **I'm overjoyed when it gives so much snow.**'

(105) *Nebenbei bemerkt: Ganz zuverlässig war die Wettervorhersage nicht, denn* ***es gab (zum Glück!) keinen Schnee****, als wir am nächsten Tag nach Hause geflogen sind.* (astrosoft.de)
'By the way: The weather forecast wasn't entirely reliable, because **it gave (luckily!) no snow** when we flew home the next day.'

(106) ***Ich habe das Gerücht gehört, dass es*** *schon bald wieder* ***Schnee geben soll!*** (mauzepow.de 'private blog')
'**I heard the rumour** that **it is supposed to give snow** again soon.'

---

[83] *apert*, Source: *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm, Neubearbeitung (1965–2018)* ('The German Dictionary by Jacob Grimm and Wilhelm Grimm, revised edn. [1965–2018]'), in the digitized version of the *DWDS* <*https://www.dwds.de/wb/dwb2/apert*>, (last accessed April 18, 2023).
[84] The blog title is a word play with *Sichtweise* 'viewpoint'. *Sicht* translates to 'view' and *Wiese* translates to 'meadow'. *Sichtwiese* is not a proper word in German.

Other than evidentiality, the last example also encompasses STANCE in the shape of an exclamation point; however, the given context does not suffice to further clarify the speaker's position on the alleged snowfall. The next section compares the frequencies and distributions of the three constructions [*es*+*schneien*], [*Schnee*+*fallen*], and [*Schnee*+*geben*] and elucidates the commonalities and differences between the usages of the three constructions.

3.2.3 The predicate construction [es+schneien] *'it+snow'*

This section surveys the *es*+PRED-Cx in combination with the precipitate *Schnee* 'snow', as in [*es*+*schneit*] 'it+snows'. The queries in the *deTenTen13* were carried out using the expressions in (xviii) and (xvii) to account for the different word orders. A total of 65,211 concordance lines where returned.

(xvii) [word="es"] [ ] {0,8} [lemma="schneien"] within <s/> (N=46,671)
(xviii) [lemma="schneien"] [ ] {0,8} [word="es"] within <s/> (N=18,540)

With respect to word order, there were more than twice as many concordance lines for PREC2 (xviii) including only one "false" return, where the concordance belonged to the other word order query. In contrast, there were 24 false returns in the PREC1 data (xvii). If this trend is true for all the PREC1 data, then approximately one-fourth of returns would have to be removed from the total frequency of 18,540 and possibly be added to PREC2. There are four figurative returns.

The false returns were removed and substituted with concordance lines matching the word order specified in the query. The following subsections present and discuss the variables itemized in Chapter 2.

3.2.3.1 The USAGE (LITERAL VS FIGURATIVE) and GENRE variables

Table 49 gives the frequencies and distribution of the literal and figurative use for the *es*+PRED-Cx and precipitate *Schnee* 'snow'. The *es*+PRED-Cx with *Schnee* 'snow' is overwhelmingly used in the literal sense (N=196).

Table 49. *Frequencies of literal and figurative use for the* es*+PRED-Cx and the precipitate* Schnee *'snow'*

USAGE / Word order	PREC1 *schneien+es*	PREC2 *es+schneien*	Total
Literal	98	98	196
Figurative	2	2	4
Total	**100**	**100**	**200**

There are only four concordance lines where the construction is applied in a metaphorical way where verb *schneien* 'snow' is being employed as a transitive verb with a direct object, such as *Gedichte* 'poems' (*es wird Gedichte vom Himmel schneien* 'it will snow poems from the sky'). Two concordance lines make use of the metaphorical *etwas/jemand schneit ins Haus*, lit. 'someone/something snows into the house', meaning someone/something arrived unexpectedly and suddenly, as in (107). Two utterances refer to digital snow snowing on websites as a winter or Christmas gimmick, as in (108). These two concordance lines were kept in the data since they simulate actual snow events.

(107) *Nach drei Jahren tropft der Wasserhahn immer noch, dafür* **schneit es** *eine Rechnung über ein paar Millionen Franken ins Haus*. (grundrechte.ch 'basic rights')
'After three years, the faucet is still dripping, but **it snows** a bill for a few million Francs into the house.'

(108) *Heute* **schneit es** *auf meinem Blog, zumindest ein paar große Flocken* ☺ . (paperartandco.blogspot.com)
'**It snows** on my blog today, at least a few big flakes ☺.

Only nine tokens overall contain negation, four of which contain positive semantic prosody, i.e., relief over there not being any snow.The distribution and frequencies of the GENRE variable for the *es*+PRED-Cx and precipitate *Schnee* 'snow' are given in Table 50. The last column shows the relative frequency of the total per 196 concordance lines. With respect to the aggregate frequencies of the GENRE variable, the corpus returns for either word order overwhelmingly belong to the WEATHER TALK genre with 163 tokens (83%), which is, incidentally, almost the same as the [*es+regnen*] construction. Second most frequent, by a wide margin, is the LITERARY type genre with 19 tokens (10%). Genre types WEATHER REPORT and METEOROLOGY have the lowest incidences with seven and four tokens respectively (3.5% and 2%). Finally, there are three literal constructions coded as OTHER, which includes instances such as the digital snow mentioned in the last subsection. Of four figurative concordance lines, one is from the LITERARY genre and three come from other sources. Regarding word order, the concordance lines are distributed almost evenly.

Table 50. *Raw and relative frequencies for the* GENRE *variable across the* es*+PRED-Cx and the precipitate* Schnee *'snow'*

Word order/ GENRE	PREC1 *schneien+es*	PREC2 *es+schneien*	Total	Relative frequency /196*
WEATHER TALK	82	81	**163**	**83%**
WEATHER REPORT	4	3	**7**	**3.5%**
METEOROLOGY	2	2	**4**	**2%**
LITERARY	8	11	**19**	**10%**
OTHER	2	1	**3**	**1.5%**
Total	**98**	**98**	**196**	**100%**
FIGURATIVE	2	2	4	divided by 200 = 2%

*literal instances only

Further scrutinizing the utterances coded as WEATHER TALK, they appear to mostly come from personal blogs where people report about their daily life and where the followers of the blog can leave comments below the blog posts. These blogs are often dedicated to special interests, such as photography, travel, or crafts, as in (109) below, where the blog owner calls herself a "yarn and knit addict" and which unashamedly expresses positive stance regarding the snow event, which in turn adds to the informality of the utterance.

(109) ***Hatte ich mir Schnee gewünscht? Jaaaaa*** *.... ich hatte!!!* ***Mein Wunsch wurde*** *auch* ***erfüllt, jeden Tag schneit es,*** *am Wochenende hörte es gar nicht auf ....* (tinas-woll-und-perlenecke.blogspot.com 'tinas-wool-and-beadcorner')
'**Did I wish for snow? Yeaaaa**....I did!!! **My wish was fulfilled,** too**, it snows every day,** and on the weekend, it didn't stop at all…'

Only few WEATHER TALK tokens originate from newspapers, reporting on sports events, or general information. The tokens coded as LITERARY almost exclusively come from fiction, only two tokens derive from from song lyrics, and none are sourced from a religious text. The four METEOROLOGY tokens relate to regional weather rather than HWW.

3.2.3.2 The TAM variable (TENSE, ASPECT, MOOD)

The distribution and frequencies of the TAM variable and its levels PAST, PRESENT (PRES with progressives marked as "/prog"), present FUTURE (pres_FUT), FUTURE (FUT) and IRREALIS (IRR) for the *es*+PRED-Cx and precipitate *Schnee* 'snow' are given in Table 51. Factors pres_FUT, FUT, and IRR are listed individually but subsumed under "future intent" in the figures (in later sections) and where they are discussed in the running text. The last column shows the relative frequency of the total per the 196 literally used returns. The distribution of the TAM variable is mostly split between concordance lines in PAST tense and PRESENT tense, which is another similarity shared

with the [*es+regnen*] construction. The use of EXISTENTIAL or GENERIC present tense amounts to 26 tokens (13%). Here, too, several (N=10) are introduced by the subordinating conjunctions *wenn* 'when/if' and *ob* 'whether' (N=1), specifying their conditionality. The "future intent" tokens have the lowest incidence (N=17, 9%). The majority of these utterances contain a modal verb (*soll* 'supposed to' N=5, *sollte* 'should/was supposed to' N=2, *kann* 'may' N=3, *könnte* 'could' N=1).

Table 51. *Raw and relative frequencies of the TAM variable within the* es*+PRED-Cx and the precipitate* Schnee *'snow' (non-literal instances removed)*

Word order / TENSE	**PREC1** *schneien+es*	**PREC2** *es+schneien*	**Total**	**Relative frequency** / 196
PAST	44	34	**78**	**40%**
PRESENT / prog	6/16= 22	9/44= 53	**75**	**38%**
present_FUTURE	10	3	**13**	**7%**
FUTURE	2	0	**2**	**1%**
IRREALIS	2	0	**2**	**1%**
EXIS_GEN	18	8	**26**	**13%**
Total	98	98	**196**	**100%**

At closer examination, both PRESENT tense and PAST tense exhibit considerable variation regarding additional TAM marking. As indicated by the slashes in the table above, out of 75 PRESENT tense tokens, a total of 59 concordance lines contain progressive aspect, i.e., roughly four-fifths. The data coded as PAST, is more varied with the results detailed in Table 52. The PAST usage of the 78 concordance lines can be further broken down into "default" *Präteritum* ('preterite/simple past', N=30), *Perfekt* (N=35), *Plusquamperfekt* ('past perfect', N=13), and progressive aspect (N=1).

Table 52. *Variations of tense and aspect within the data comprehensively coded as PAST for the* es*+PRED-Cx and the precipitate* Schnee *'snow'*

Word order/ PAST	**PREC1** *schneien+es*	**PREC2** *es+schneien*	**Total**
Preterite	18	11	**29**
Perfekt	11	24	**35**
Past perfect	4	9	**13**
Progressive	1	0	**1**
Total	34	44	**78**

Perfekt is known as the "conversational past" in German and it is easy to see the differences in formality between the use of *Perfekt* in (110) as opposed to the preterite in (111), which is mostly used for more formal texts such as reports, demonstrated below. The first example clearly addresses the reader of the post or the message on this personal blog and reports on the weather conditions at the speaker's location. It exhibits stance, expressing not only the personal

positive opinion about the weather conditions (*herrlich winterlich* 'delightfully wintery'), but also takes into account that many of the readers probably do not share the speaker's opinion. The second example stems from an online newspaper reporting on some historical weather conditions. Still, the latter is strictly speaking not a WEATHER REPORT and both examples were coded as WEATHER TALK (as are most of the examples in PAST tense. There are only four PAST concordance lines coded as LITERARY and they are all in preterite.) The concordance lines in past perfect, too, are mainly conversational.

(110) ***Heute hat es bei uns geschneit*** *und draußen ist es nun herrlich winterlich. (Ich weiss, viele von euch sehnen sich schon jetzt nach dem Frühling…)* (mamaskram.blogspot.co.at 'mama's-stuff.blogspot')
'**Today it has snowed here** and it is now delightfully wintery outside. (I know, many of you are already longing for spring).'

(111) *Am eindeutigsten hat sich diese Kälteperiode von 1550 bis 1700 bemerkbar gemacht. Vorboten gab es aber schon früher zuhauf.* ***1318 schneite es am 30. Juni in Köln****, und die Jahre von 1345 bis 1347 sind als "Jahre ohne Sommer" in die Annalen eingegangen.* (woz.ch, *Die Wochenzeitung* – 'The Weekly Newspaper')
'This cold period made itself felt most clearly from 1550 to 1700. But there were plenty of harbingers earlier. **In 1318 it snowed in Cologne on June 30**, and the years from 1345 to 1347 have gone down in the annals as "years without a summer".'

Lastly, the use of quasi-auxiliaries is not as pronounced as for the *es*+*regnen*-Cx although the focus on the incipient phase of the ME remains: *anfangen/beginnen* 'start/begin' occur with 10 tokens and *aufhören* 'stop' occurs four times. Two utterances are "other" (*nicht mehr* 'not anymore' and *schon wieder nicht mehr* 'yet again not anymore'). All of the quasi-auxiliaries are inflected for some kind of PAST tense, mostly *Perfekt*.

3.2.3.3 The SPATIAL modifier variable

Table 53 shows the distribution and frequencies of different SPATIAL modifiers occurring within the corpus returns for the *es*+PRED-Cx and precipitate *Schnee* 'snow'. Slightly more than half of the returns do not contain any type of SPATIAL modification whatsoever (ZERO). Where there is a SPATIAL mention, it tends to be a LOCATION; roughly 10% of the LOCATIONS do not directly modify the ME but are given in the wider context. The small remainder of 1.5% is made up of two GOALS and one PATH, where in one utterance the GOAL and the PATH are combined: *bis ins Tal herunter* 'down into the valley'. In the other GOAL utterance (*bis in tiefe Lagen* 'to low

locations/altitudes'), the path is not additionally expressed through a surface form like *herunter* 'down'. There are no SOURCES. The semantic make-up of the LOCATIONS is discussed below.

***Table 53**. Raw and relative frequencies of the* es*+PRED-Cx for the SPATIAL modifier variable and the precipitate* Schnee *'snow'*

Word order/ SPATIAL modifier	**PREC2** *es+schneien*	**PREC1** *schneien+es*	**Total**	**Relative frequency** / 196
ZERO	39	62	**101**	**51.5%**
LOCATION	59	35	**94**	**48%**
GOAL	1*	1*	**2**	**1%**
PATH	1*	0	**1**	**0.5%**
SOURCE	0	0	**0**	**0%**
Total	**100**	**98**	**198**	**101%**
COMBINATIONS	1	0	1	0.5%

**bis in(s)* lit. 'until into', *herunter* 'down/downwards' (i.e., to a lower position/place)

Modifiers of LOCATION let themselves easily be grouped into a small number of semantic categories despite their relatively high frequency. The four categories are "deictics", "location", "cardinal direction", and "other". Especially the largest category of "deictics" (N=43) consists of basically five types: *hier* 'here' (N=15), *draußen* 'outside' (N=12), *bei uns/euch* 'at our/your place' or 'here' (N=11), and *dort* 'there' and *gebietsweise* 'in some areas' (N=2 each). The category "location" mostly contains mentions of cities/countries (e.g., *in Berlin/in Deutschland* 'in Berlin/in Germany', N=13) as well as regions or mountains (eight tokens, e.g., *in Ostwestfalen* 'in East Westphalia', *in den Anden* 'in the Andes') and very specific, "focused" locations, e.g., *am Ausgang des Felberntauerntunnels* 'at the exit of the Tauern Road Tunnel'. The latter obviously does not mean that it snowed only in that location exactly, but, presumably, rather, that it was not snowing at the time the speaker entered the tunnel, which is not unthinkable considering that the tunnel in question has a length of more than 6.5 kilometers. There are five spatial mentions which contain "cardinal directions", such as *im Norden der USA* 'in the north of the US'. Lastly, the group "other" (five tokens) contains another temporal modifier, which combines with a LOCATION: *bei der Ankunft im Vogelsberg* 'upon arrival in Vogelsberg' as well as a couple of phrases referring to altitude (*bis in tiefe Lagen/in tieferen Lagen* 'in[to] lower altitudes', i.e., the two goals and the path. There are conspicuously few combinations, such as *im Norden, in Jukkasjärvi* 'in the north, in Jukkasjärvi'.

3.2.3.4 The NON-SPATIAL modifier variable (TEMPORAL, QUANTITY, QUALITY, STANCE)

The distribution and frequencies of the NON-SPATIAL modifier variable for the *es*+PRED-Cx and precipitate *Schnee* 'snow' are given in Table 54. The last column shows the relative frequency of

the total per 196 concordance lines with literal meaning. SPATIAL mentions are listed in the last row and count towards COMBINATIONS as well as the concordance lines that contain ZERO modification. The most frequent modifier is TEMPORAL with more than two-thirds of the concordance lines containing at least one modifier of TIME. The second most frequent modifier is STANCE; in almost half of the data (N=91; 46%), the speaker expresses some sort of attitude towards the ME. Modifiers that express the QUALITY or MANNER of the snow falling are not quite half as frequent as expressions of STANCE. The lowest incidence is provided by modifiers of QUANTITY (N=30; 15%). These instances include and reduplication (e.g., es *schneit und schneit* 'it snows and snows', N=5) as well as negation, which specifies ZERO quantity (N=9). Only four percent of the *es+schneien* 'it+snows' returns do not contain any modifier whatsoever, but almost two-thirds of the returns (N=129; 64.5%) combine at least two different modifiers (this includes SPATIAL mentions).

Table 54. *Raw and relative frequencies of the NON-SPATIAL MODIFIER variable (TEMP, QUANT, QUAL, STANCE) within the* es*+PRED-Cx and the precipitate* Schnee *'snow' (literal use only)*

Word order/ MODIFIER	**PREC1** *schneien+es*	**PREC2** *es+schneien*	**Total**	**Relative frequency** /196
TEMPORAL	72	61	133	**68%**
QUANTITY	14	16	30	**15%**
QUALITY	30	13	43	**22%**
STANCE	39	52	91	**46%**
Total	**155**	**142**	**297**	**199%**
ZERO	2	6	8	4%
COMBINATIONS	73	56	129	66%
SPATIAL modifier	52	30	82	42%

TEMPORAL modifiers are also the most frequent stand-alone modifiers (N=26), as illustrated in (112), followed by STANCE (N=16), as illustrated in (113).

(112) ***Es hat die Nacht geschneit****, es sind gerade knapp -3°C und es liegen etwa 8cm Schnee.* (uni-dortmund.de 'University Dortmund')
'**It has snowed overnight**, it is barely -3°C and there are about 8cm of snow on the ground.'

(113) ***Es ist eine wunderbare Zeit, es schneit****, überall duftet es nach Weihnachtsplätzchen und Glühwein.* (artikelwand.de 'article wall')
'**It is a wonderful time, it snows**, it smells like Christmas cookies and mulled wine everywhere.'

The following subsections elaborate on the nature of each type of modifier individually starting with modifiers of TIME.

Modifiers of TIME. Modifiers of TIME are not just the most frequent but also encompass the highest number of semantic categories and individual types with quite a lot of overlap between categories. Out of the ten categories having at least five tokens each, the one with the most tokens (but not necessarily types) are "deictics" and "frequency" with more than 50 tokens altogether. The most frequent modifier types are adverbs *wieder* 'again' (N=24, including all combinations with other TEMPORAL modifiers) and *heute* 'today' (N=16, including all combinations with other TEMPORAL modifiers). This also makes adverbs the most frequent syntactic PoS. Other frequent deictic types are *jetzt* 'now' (N=6) and its synonyms *im Moment*/*momentan*/*gerade* 'at the moment/now', which furthermore function as progressive markers. "Frequency", moreover, contains items such as *oft* 'often', *manchmal* 'sometimes', *hin und wieder* 'now and then', *vereinzelt* 'sporadically', or *an zwei von drei Tagen* 'on two out of three days'.

The category "time of day" is as frequent as modifier heute 'today' by itself (N=16, including a few simple combinations such as *heute Nacht wieder,* lit. 'today night again', i.e., 'last night again'). The time of day which is particularly frequent is *Nacht* 'night' (N=10) or variations thereof, such as *über Nacht* 'overnight', *in der Nacht* 'during the night' or *die ganze Nacht* 'all night'. Obviously, these could also be analyzed as timespans. The item *Nacht* 'night' may also indicate that the speaker did not actually witness the ME itself but rather the result, i.e., they woke up to find that there is snow outside while still referring to the event rather than the result, i.e., the quantity of snow on the ground, as in (112), as indicated by the low frequency of QUANTITY modifiers. Most "time of day" modifiers are nominal, but there are also a couple of adverbs, such as *tagsüber* 'during the day' and *morgens* 'in the morning'.

The category "combinations" (N=25) ranges from more simple combinations with two items, like *heute … wieder* 'today … again', which combines deixis and frequency, to more complex, like *Freitag, 12. Februar 2010. Seit dem letzten Beitrag … schon wieder* 'Friday, 12 February 2010. Since the last post … yet again', which combines the day of the week with the exact date, a point in time, and frequency. However, the simpler combinations of no more than two tokens from each one category prevail. "Timespans", such as *seit Tagen* 'for days', often overlap with "time/part of the day", (see above, *die ganze Nacht* 'all night') but also specify the starting point and/or that the snow event is ongoing (e.g., *seit gestern fast kontinuierlich* 'almost continually since yesterday' and *seit Mittwoch …. fast ununterbrochen* 'since Wednesday … almost uninterruptedly').

The categories “season” and “month” (N=14 combined) predominantly contain instances of *Winter* ‘winter’ (N=9). The remainder are different months (except for *November*, which occurs twice). Like “month”, “day of week” (N=5) and “exact date” (N=7) are low frequency categories and the latter often appear in combinations (see above). Lastly, the category “other” contains items of “sequentiality”, such as *erst*/*anfangs* ‘at first’, but also idiosyncratic instances such as *autozufahren, während* ‘while driving’, *bei der Ankunft im Vogelsberg mal wieder* ‘upon arrival in Vogelsberg again’ and *auf der Rückfahrt/unterwegs* ‘on our way [back]’.

Modifiers of QUANTITY. As with the *Regen* ‘rain’ precipitate in the *es*+PRED-Cx (N=18; 9%), the QUANTITY modifier for the *Schnee* ‘snow’ precipitate is the least frequent (N=30; 15%). Despite the overall low incidence, the modifiers can be grouped into five categories. The quantities that modify the ME are either “small quantities” (N=6), such as *geringfügig* ‘trifling’, *wenig* ‘little’, or *etwas* ‘some/a little’ or “larger quantities” (N=6), such as *reichlich* ‘plenty’ and *so viel* ‘so much’ or some fluctuation as in *mal mehr mal weniger* ‘sometimes more sometimes less’. The third category is “reduplication” (N=5), where the verb (but not necessarily the argument) is repeated either two or three times, as in *es schneit und schneit und schneit* ‘it snows and snows and snows’. Alternatively, reduplication could be grouped with the “large(r) quantities” category. Obviously, these instances of reduplication may also include an element of time. Additionally, there are three resultative tokens, where the quantity that it snowed is specified in centimeters; twice, these are large quantities (*20cm*, *40cm*), and once a comparatively smaller quantity (*5cm*). The remainder are negated concordance lines, where the quantity is ZERO (N=9).

Modifiers of QUALITY. Modifiers of QUALITY or MANNER, despite being relatively infrequent with 22%, can be grouped into five categories, including the catch-all “other”. Syntactically, the most frequent are adverbial modifiers, which express the “intensity or severity” of the snow falling. These instances could obviously also be analyzed as instances of QUANTITY. There are 14 concordance lines containing mostly adverbs such as *heftig* ‘severely’ and *kräftig* ‘heavily’ (each three tokens) as well as more informal idioms, such as *wie verrückt*, *wie wild*, and *wie Bolle* (all three meaning ‘like crazy’, i.e., to a great extent). Furthermore, there is one consecutive clause, consisting of the slightly altered turn of phrase *dass man keinen Hund vor die Tür jagen möchte* ‘that you would not chase a dog out the door’ (the original expression is *Bei diesem Wetter jagt man keinen Hund vor die Tür* ‘In this weather you would not turn a dog out’ lit. ‘In this weather

you do not chase a dog out the door'). Nine utterances pertain to the "continuity" or "incessancy" of the event; these are, for example, adjectives like *unaufhörlich* 'incessantly', *ununterbrochen* 'without interruption', and *kontinuierlich* 'continually', *ohne Pause* 'without pause', or the separate clause *es hört gar nicht mehr auf* 'it won't stop'. The category describing the "insubstantiality" of the snow has a fairly low incidence (N=6) of the very same type, i.e., *leicht* 'lightly' or *ganz leicht* 'very lightly'. Five concordance lines are transitive where the object is modified and describes the quality of the snowflakes such as *schöne, grosse, dicke Flocken* 'nice, big, thick flakes', *dicke Flocken* 'thick flakes', and *große Flocken* 'big flakes'. Additionally, there is an example where it snows *Schneeflöckchen* but then develops into more severe weather conditions, as in (114).

(114) *Seit ein paar Tagen* ***schneit es*** *immer wieder,* ***von Schneeflöckchen zu Schneefall, dann ein ausgewachsener Schneesturm.*** (eden-borkum.de 'granny au pair from island Borkum)
'**It's been snowing** for a few days again and again, **from snowflake to snowfall, then a full blown blizzard**.'

Take note that only *Schneeflöckchen* can be the object of *schneit es* 'it snows' since **Es schneit einen Schneefall/Schneesturm* *'It snows a snowfall/snow storm/blizzard' is odd at best, if not ungrammatical. Lastly, the category "other" (N=9) is a bit of a potpourri. It includes items of MANNER, such as *in dichten Flocken* 'in dense flakes', *still und leise vor sich hin* 'softly and quietly away', *dass der Schnee liegenbleibt* '[in a manner] that the snow stays', *plötzlich* 'suddenly' but also *zum ersten Mal* 'for the first time' (N=3). The latter modifier is "snow-specific" as speakers usually do not make mention of the first rain, sleet, or hail – unless they specifically refer to the "first rain of the year" in a new year or possibly in a rain-deprived region. However, even then, it appears that further modification would be required, as in *the first rain since*… or *the first rain in months*. In contrast, in countries that have all four seasons, the *first snow* usually refers to the first snow of the winter season.

Modifiers of STANCE. Modifiers of STANCE are very frequent and represent the second-highest category of NON-SPATIAL modification of the *es+schneien* returns with 46%. They are fairly evenly distributed between positive (N=29) and negative (N=32) semantic prosody. Additionally, they may express something "other", like surprise or indifference. The majority of stancetaking here leans more toward the negative. In addition to the clearly negative items, there are an additional eight subordinate clauses introduced by *selbst/auch wenn* 'even if', (*egal) ob* '(no matter) whether', which also lean more towards the negative. Overall, negativity is conveyed less emphatically than positive attitude by expressions such as *glücklicherweise nicht*

'fortunately not' or *leider* 'unfortunately' but also by communicating (possible) negative consequences, such as *kam zu einem Ausfall der Ernte und in der Folge zu Hungersnöten* 'resulted in crop failure and, as a consequence, famine', *haben gerade noch so nach Hause geschafft, bevor vor uns ein Unfall geschah* 'just so made it home before an accident happened', *hab damit null gerechnet und wäre auf dem Weg zur Arbeit fast erfroren* 'didn't expect it at all and almost froze to death on my way to work'. Other more personal, and more forceful expressions of disgust include interjections like *Bääh!* 'Yuck!' and *bbbbbrrrrrrr!!!!,* adjectives like *ekelhaft* 'disgusting' and *geärgert, dass es schon wieder …* 'irritated that yet again it…'. Also, confirm example (115), in which the speaker forthrightly describes their negative emotional state at witnessing the weather outside.

(115) ***Ich könnte weinen****, wenn ich aus dem Fenster schaue.* ***Am WE war so tolles Wetter und jetzt schneit es wieder. Sch..... Ich dreh bald durch****. Seit Oktober haben wir Schnee, und das nicht wenig. Wo bleibt der Frühling?* (bcm.de 'Precon diet chat forum')
'**I could cry** when I'm looking out the window. **The weather was so awesome on the weekend and now it's snowing again. Sh…. I'm about to lose it.** We've been having snow since October and not too little. Where's spring?'

A positive attitude towards snow appears to be phrased overall more vigourously as documented by the following expressions: *toll* 'awesome', *das Beste daran* 'the best part', *endlich* 'finally', *herrlich winterlich* 'wonderfully wintry', *eine wunderbare Zeit* 'a wonderful time', s*o bin ich nun geradezu verzückt* 'so now I'm downright ecstatic' or people literally claiming that they (and/or their four-legged friends) love the weather situation: *ich/Goliath/Bella liebe/liebt es* 'I/Goliath/Bella love it'.

Of course, there are also more elliptical utterances, such as *Dieses Jahr weiße Weihnachten? Sieht gut aus*. 'This year white Christmas? Looking good.' and *Ich warte drauf, aber wer weiss, wann das bei uns etwas wird* 'I'm waiting for it, but who knows when it's going to happen here'. A few expressions categorized as "other" (N=19) convey positivity and negativity at the same time. They do so by taking different views into account, as in (116), where the perspective is that of the family dogs who love snow, as opposed to their owners, who don't (cf. also (110) in Section 3.2.3.2above).

(116) ***WIR lieben ja den Schnee****, aber die Zweibeiner haben so langsam wirklich genug*.
(schlappohren.over-blog.de 'floppy-ears.over-blog')
'**WE love the snow**, but the bipeds are getting really tired of it.'

Lastly, there are "other" concordance lines which are meant ironically: *Schnee! Im Winter! Hammer* 'snow! In Winter! Wow', *kaum zu glauben* 'hard to believe', especially mocking public transport being "surprised" by snowfall every winter again, and occasions of anticipation, which could go either way: *ich bin mal gespannt* 'I'm curious/keen/anxious' or simply an exclamation like *Aaaaaaaaaahhhhhhhh!.* This paragraph concludes the section on the *es*+PRED-Cx. The next section details the distribution and frequencies of variables across the ARG+*fallen* corpus returns and the *Schnee* 'snow' precipitate.

*3.2.4 Distribution and frequencies of variables across the three constructions (ARG+*fallen*, ARG+*geben*, es+PRED) with the precipitate* Schnee *'snow'*

In this section, the frequencies and distribution of variables by construction, i.e., the *es*+PRED-Cx and the two ARG-Cxx in combination with the *Schnee* 'snow' precipitate, as surveyed up to this point, are compared against one another. Despite the similarities, the constructions appear to have carved out their niches with respect to certain variables in a similar manner to the *Regen* 'rain' precipitate constructions as already foreshadowed in the introductory paragraph of this section. In addition to the commonalities and differences highlighted there, more subtle differences have crystallized in the semantics of the variables, such as the positive semantic prosody within the [*Schnee+geben*] utterances marked for STANCE as opposed to the negative semantic prosody associated with the [*es+schneien*] construction in contrast to the overall infrequent expression of STANCE within the [*Schnee+fallen*] data sets. The following sub-section starts with the literal and figurative USAGE and GENRE variables, where the former separates the ARG-Cxx from the *es*+PRED-Cx and the latter shows an overwhelming use of the WEATHER TALK genre type, which, overall, turns out to be even more frequent than for the *Regen* 'rain' precipitate data sets.

3.2.4.1 The USAGE (LITERAL vs FIGURATIVE) and GENRE variables across the three constructions

The distributions and frequencies of literal and figurative use are given in Table 55. The three constructions are overwhelmingly used in a literal sense. The two ARG-Cxx contain no instances of figurative uses, only the [*es+schneien*] construction has four instances of figurative use.

Table 55. *Aggregate frequencies of literal and figurative* USE *by construction type (ARG+*fallen*, ARG+*geben*, and* es*+PRED) with the precipitate* Schnee *'snow'*

Construction / USAGE	*Schnee+fallen*	*Schnee+geben*	*es+schneien*	Total
LITERAL	200	200	196	**596**
FIGURATIVE	0	0	4	**4**
Total	**200**	**200**	**200**	**600**

The distribution and frequencies of construction by GENRE are given in Table 56. The relative frequency per 596 literally used concordance lines is given in the last column. The absolute frequencies are furthermore displayed in Figure 14. The frequencies summarize what has been established individually: WEATHER TALK is the predominant level of GENRE across all three constructions. Almost one-third of the data are coded as WEATHER TALK. The second highest incidences, not even one-tenths as frequent as WEATHER TALK, but evenly distributed, are the genre types METEOROLOGY and LITERARY. The lowest incidence of genre type is comprised of WEATHER REPORTS with roughly one-tenth of the WEATHER TALK returns. The remainder (OTHER, N=1 and FIGURATIVE, N=4) amount to roughly 3% of the returns and were removed from Figure 14 for reasons of simplification.

Table 56. *Raw and relative frequencies of the* GENRE *variable across the three constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED) in combination with the precipitate* Schnee *'snow'*

Construction / GENRE	*Schnee+fallen*	*Schnee+geben*	*es+schneien*	Total	Relative Frequency /596
WEATHER TALK	112	163	163	438	73%
WEATHER REPORT	32	6	7	45	7.5%
METEOROLOGY	32	18	4	54	9%
LITERARY	22	13	19	54	9%
OTHER	2	0	3	5	0.8%
Total	200	200	200	600	99.3%
FIGURATIVE	0	0	4	4	(divided by 200) 2%

Pearson's chi-squared test: 64.816, df = 6, p-value = 4.703e-12

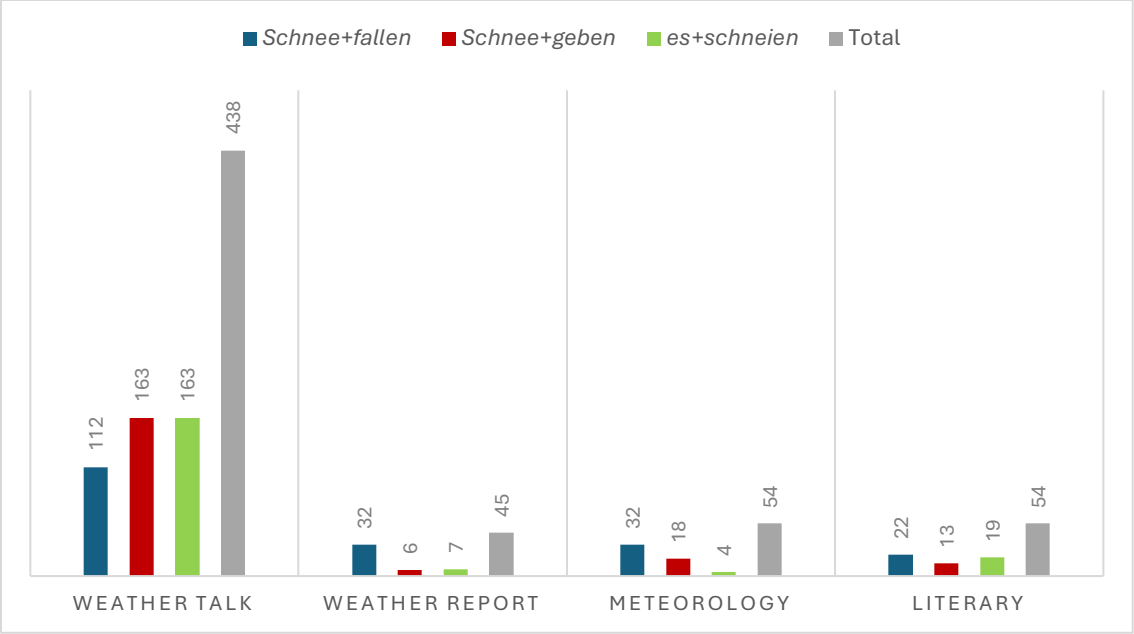

Figure 14. *Absolute frequencies of the* GENRE *variable across the three constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED) in combination with the precipitate* Schnee *'snow' ("other" and figurative instances removed)*

Figure 15 shows the contributions of each construction to each type of GENRE in percentages. The most noticeable findings are that genre type WEATHER TALK is the most evenly distributed between the three construction types. At the same time, the [*Schnee+fallen*] construction contributes the least to genre type WEATHER TALK with about one-quarter of all WEATHER TALK data points and contributes the highest proportion of every other type of GENRE, especially WEATHER REPORTS and METEOROLOGY. It contributes almost twice as much to the genre type METEOROLOGY than the [*Schnee+geben*] construction, while the [*es+schneien*] construction makes up the meagre remainder of 7%. Lastly, while the [*Schnee+geben*] construction and the [*es+schneien*] construction contribute exactly evenly to the WEATHER TALK genre, and fairly evenly to genre type WEATHER REPORTS, the [*es+schneien*] construction is closer to the [*Schnee+fallen*] construction regarding the LITERARY genre. Here, the [*Schnee+geben*] construction only accounts for roughly one quarter or the returns.

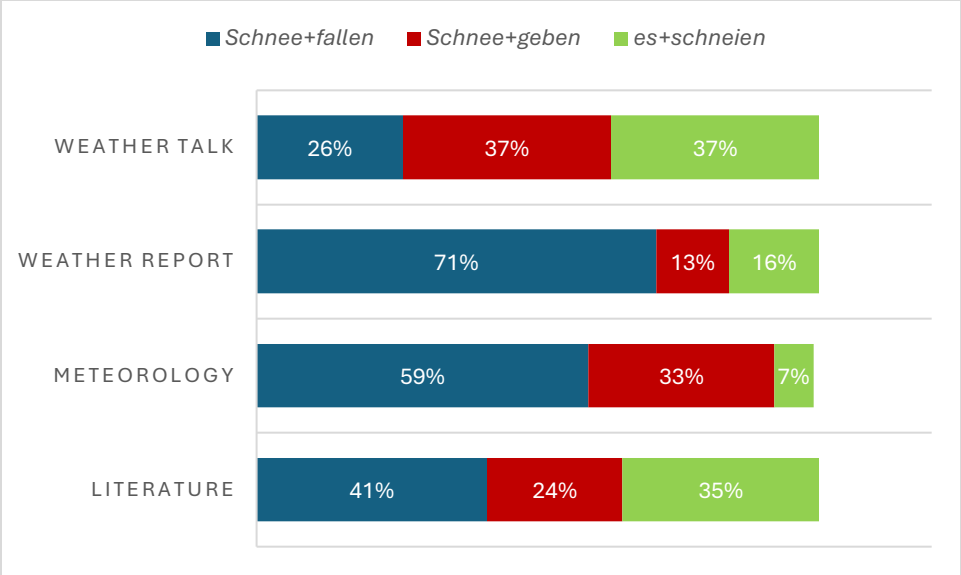

Figure 15. *Relative frequencies in percentages of the* GENRE *variable across constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED) with the precipitate* Schnee *'snow'*

In summary, the [*Schnee+fallen*] construction is the most equally (relatively) distributed across types of genre, while WEATHER TALK and LITERARY are the types of genre which attract the three constructions most proportionately (relatively). WEATHER REPORTS, in particular, but also genre type METEOROLOGY, most clearly attract the [*Schnee+fallen*] construction. This indicates that the [*Schnee+fallen*] construction is used in a more formal manner than the other two constructions since it occurs in more formal genre types.

3.2.4.2 The TAM (TENSE, ASPECT, MOOD) variable across the three constructions

This section deals with the distributions and frequencies of the TAM variable across the three constructions [*Schnee+fallen*], [*Schnee+geben*], and [*es+schneien*], as given in Table 57. The absolute frequencies are furthermore displayed in Figure 16. The distribution of the TAM variable for the precipitate *Schnee* 'snow' across constructions is not homogenous. Overall, just like we have seen for the *Regen* 'rain' precipitate, *Schnee*-MEs are most likely to occur in the PAST tense. This is true for more than one third of all concordance lines. Second most frequent, with about two-thirds as many instances, are tokens of "future intent", i.e., present_FUTURE, FUTURE, and IRREALIS. This puts them even ahead of PRESENT tense, by about two percentage points. Considered individually, it is striking how little the [*es+schneien*] construction attracts the periphrastic future tense. The few instances found in the returns are chiefly reserved for the two

ARG-Cxx. The EXISTENTIAL/GENERIC use is predominantly attracted by the [*Schnee+fallen*] construction, which accounts for roughly half of the returns, the other half is evenly distributed between the [*Schnee+geben*] and the [*es+schneien*] constructions.

Table 57. *Raw and relative frequencies of the TAM variable by construction type (ARG+*fallen*, ARG+*geben*, and* es*+PRED) with the precipitate* Schnee *'snow' (FIGURATIVE/OTHER removed)*

Construction / TAM factor	***Schnee+fallen***	***Schnee+geben***	***es+schneien***	**Total**	**Relative Frequency** /596
PAST	54	86	78	**218**	**37%**
PRESENT	24	32	75	**131**	**22%**
pres_FUT	50	33	13	**96**	**16%**
FUTURE	16	14	2	**32**	**5%**
IRREALIS	6	9	2	**17**	**3%**
EXIS_GEN	50	26	26	**102**	**17%**
Total	**200**	**200**	**196**	**596**	**100%**

Pearson's chi-squared test: 90.193, df = 10, p-value = 4.903e-15

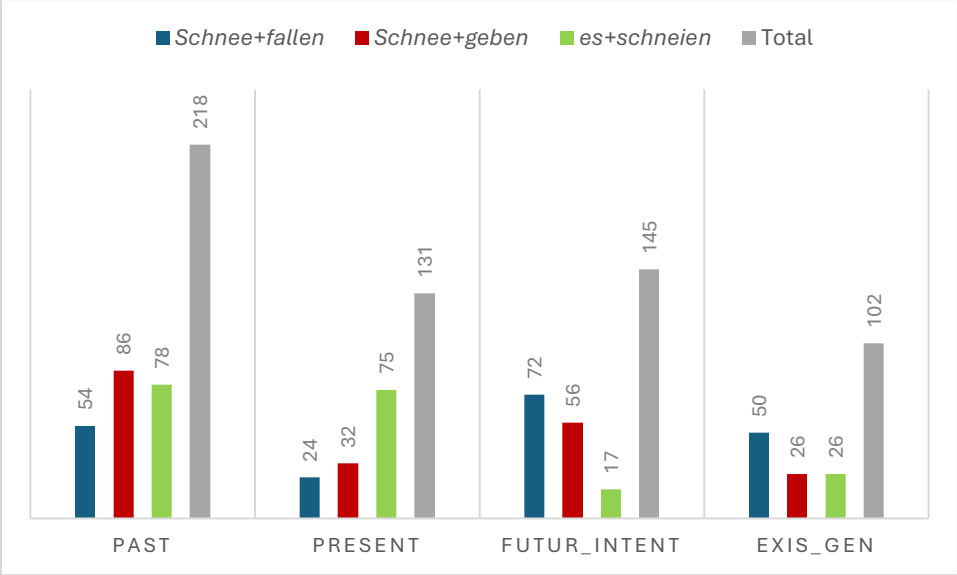

Figure 16. *Absolute frequencies of construction (ARG+*fallen*, ARG+*geben*,* es*+PRED) by TAM variable in combination with the precipitate* Schnee *'snow' (levels pres_FUT, FUT, and IRR subsumed under "FUTURE_intent"; FIGURATIVE/OTHER removed)*

If we look at the distribution of the TAM variable by construction –the relative percentual proportion that each construction contributes to the overall frequency of each TAM level is given in Figure 17– we find a similarly skewed distribution, which reveals how the different constructions tend to favour one TAM level over another. At first glance, the [*Schnee+fallen*] construction appears to strongly favour utterances in EXISTENTIAL/GENERAL use or to express future intent, where it contributes half of the data of either TAM level. As with precipitate *Regen* 'rain', we (incorrectly) may have expected the [*Schnee+geben*] construction to prevail here as

the "designated" existential construction. The [*es+schneien*] construction contributes the lowest proportion of future intent (12%), but the highest percentage of PRESENT tense, which goes hand-in-hand with its use where the speaker seemingly immediately and directly experiences the ME. For PRESENT tense and future intent, the numbers are almost inverted for the [*Schnee+fallen*] construction and the [*es+schneien*] construction. The [*Schnee+geben*] construction appears to be the most balanced construction of the three, but it makes up the highest proportion of PAST tense and also contributes a fairly high percentage of the future intent utterances. Its lowest contribution to a TAM level is that of PRESENT tense where it furthermore displays a noticeable absence of present progressive (N=1). In contrast, there is a noticeable presence of progressive marking within the [*es+schneien*] returns (N=60 out of N=75). The [*Schnee+fallen*] returns also contain a fairly large number of progressive marking (N=17) albeit within a much smaller set (N=24) overall.

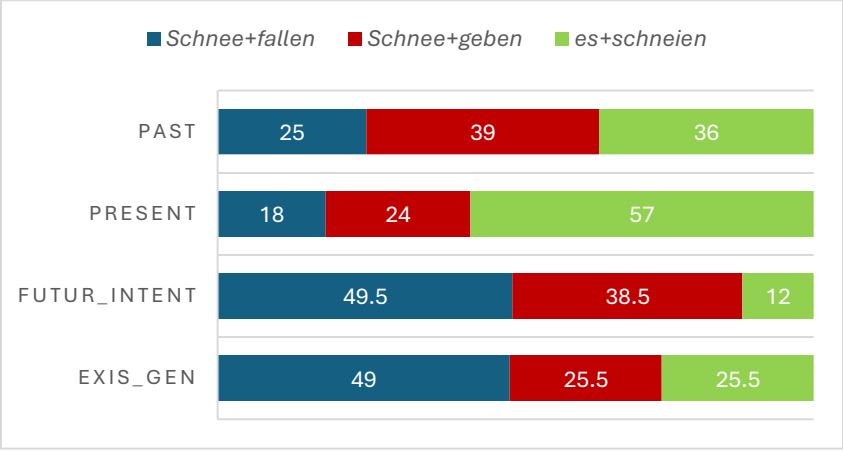

Figure 17. *Relative contribution of each construction (ARG+*fallen*, ARG+*geben*, es+PRED) to level of* TAM *for the* Schnee *'snow' precipitate data in percentages*[85]

The [*es+schneien*] construction is furthermore the one, which exhibits the most diversity within all PAST utterances encompassing preterite, *Perfekt*, past perfect, and progressive aspect. By way of contrast, [*Schnee+fallen*] utterances are exclusively inflected for the preterite. [*Schnee+geben*], on the other hand, contains several utterances in *Perfekt* as well as one utterance in past perfect. Keep in mind that *Perfekt* is known to be the "conversational past" and, thus, is presumably less formal. In a similar manner to progressive marking, the [*es+schneien*] construction is the most likely to attract quasi-auxiliaries like 'start' and 'stop'

[85] The percentages were arrived at by dividing the total per TAM level of each construction by the total of all three constructions per TAM level combined.

(N=16). [*Schnee+fallen*] contains a few instances of quasi-auxiliaries (N=6) and [*Schnee+geben*] does not contain a single quasi-auxiliary. Within the quasi-auxiliaries, there is a clear focus on the incipient phase of the event over focus on the terminative phase. The findings on additional TAM marking lead to the conclusion that both [*Schnee+geben*] and especially the [*es+schneien*] construction are much more immediate and entangled with the experience of the ME than the [*Schnee+fallen*] construction. The reverse is true for the [*Schnee+fallen*] construction attracting future intent and EXISTENTIAL/GENERIC use. Lastly, with respect to conditionality, the [*es+schneien*] construction contains 15 instances of MEs in subordinate clauses introduced by *wenn* 'when/if' and five tokens of *ob* 'whether/if' while the ARG-Cxx only contain 11 tokens (*Schnee+geben*) and five tokens respectively (*Schnee+fallen*) of subordinate clauses introduced by *wenn* and *ob*. Obviously, these clauses do not refer to actualized events as such but move within the realm of irrealis. At the same time, they point to situations that usually come true under a given condition, such as snowboarders/skiers claiming to always be happy when it snows or it being unfavourable running conditions when it snows. Accordingly, they were coded as EXISTENTIAL/GENERIC rather than IRREALIS, which may contribute to the lower "future intent" proportion of the constructions in question.

3.2.4.3 The SPATIAL modifier variable across the three constructions

This section discusses with the distributions and frequencies of the SPATIAL modifier variable across the three constructions [*es+schneien*], [*Schnee+fallen*], and [*Schnee+geben*]. The SPATIAL modifiers are LOCATION, GOAL, SOURCE, PATH, and absence of SPATIAL modification (ZERO). The distributions and frequencies of SPATIAL modifier by construction are given in Table 58 and the absolute frequencies are visualized in Figure 18. Overall, LOCATIONS are the most frequent SPATIAL modifier with 54%. This includes LOCATIONS that do not immediately modify the ME and are present in the peripheral context. These are predominantly present within the [*Schnee+geben*] returns. Only if LOCATIONS are restricted to having to directly modify the ME within the confines of the same sentence, do ZERO paths outnumber all other SPATIAL modifiers (marginally) with one percentage point more than LOCATIONS (42% vs 41%).[86] Regarding the more peripheral and less frequent SPATIAL modifiers GOAL, SOURCE, and PATH, GOALS are the most frequent (N=15), followed by SOURCES (N=9) and then PATHS (N=4). Together, these make up only 5% of the data and are

[86] These numbers were arrived at separately and are not given in Table 58.

mostly attracted by the [*Schnee+fallen*] construction (N=25), a tendency equally manifest in the [*Regen+fallen*] returns. The [*Schnee+geben*] construction contains not a single one of these peripheral SPATIAL modifiers.

Table 58. *Overall distribution and absolute frequencies of the SPATIAL modifier variable within each individual construction (ARG+*fallen*, ARG+*geben*, and* es*+PRED) with the precipitate* Schnee *'snow'*

Construction / SPATIAL modifier	***Schnee+fallen***	***Schnee+geben***	***es+schneien***	**Total**	**Relative Frequency** /596
LOCATION	122	108	94	**324**	**54%**
GOAL	13	0	2	**15**	**2.5%**
SOURCE	9	0	0	**9**	**1.5%**
PATH	3	0	1	**4**	**0.7%**
ZERO	59	92	101	**252**	**42%**
Total	**206**	**200**	**198**	**604**	**100.7%**
COMBINATIONS	5	0	1	6	6

Pearson's chi-squared test: X-squared = 54.453, df = 4, p-value = 4.23e-11

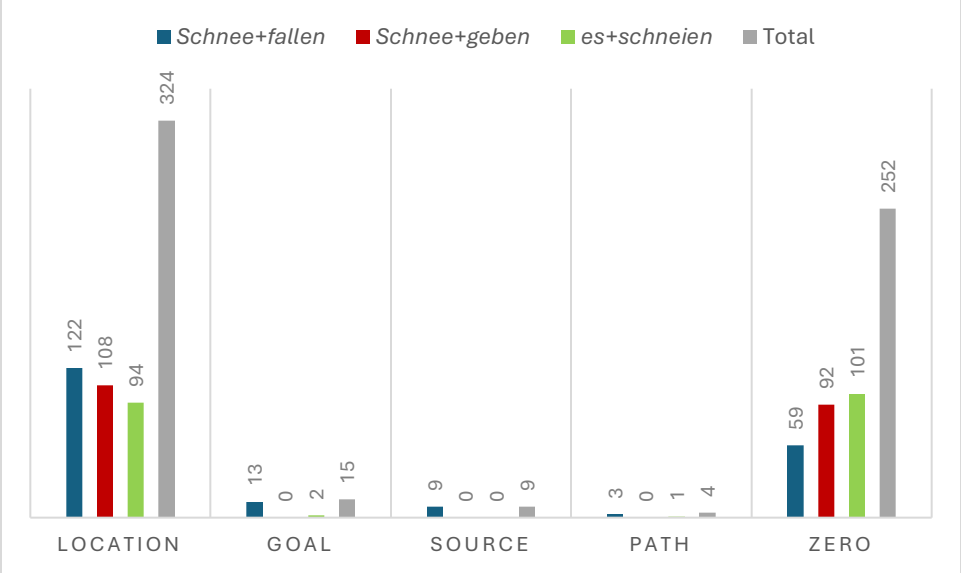

Figure 18*. Overall distribution and frequencies of SPATIAL modifiers across the three constructions (ARG+*fallen*, ARG+*geben*, and* es*+PRED) with the precipitate* Schnee *'snow'*

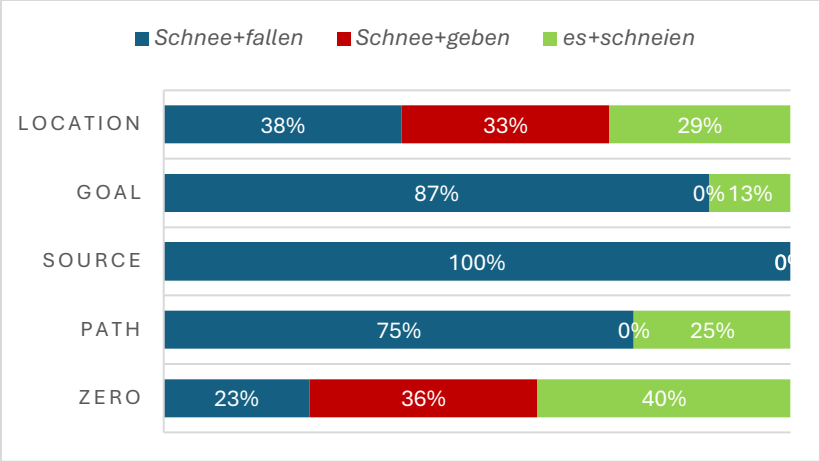

Figure 19. *Relative contribution of each construction (ARG+*fallen*, ARG+*geben*, es+PRED) to each* SPATIAL *modifier type for the* Schnee *'snow' precipitate data in percentages*[87]

In summary, the [*es+schneien*] construction contains the highest incidence of ZERO SPATIAL modification. Both the [*es+schneien*] construction and the [*Schnee+geben*] construction predominantly contain LOCATIONS and ZERO modification, and [*Schnee+fallen*] contains the majority of all SPATIAL modifiers of either type.

3.2.4.4 The NON-SPATIAL modifier variable across the three constructions

This section examines the distributions and frequencies of the NON-SPATIAL modifier variable across the three constructions [*es+schneien*], [*Schnee+fallen*], and [*Schnee+geben*]. The NON-SPATIAL modifiers are TEMPORAL, QUANTITY, QUALITY, and STANCE. Table 59 gives the distributions and frequencies of different types of NON-SPATIAL modifier by construction. SPATIAL modifiers are given in the last row and contribute to the rows COMBINATIONS and ZERO. That is to say, a concordance line that encompasses a LOCATIVE and a TEMPORAL modifier contains a COMBINATION of modifiers and any concordance line that encompasses neither a SPATIAL nor a NON-SPATIAL modifier is coded as ZERO. The absolute frequencies are visualized by Figure 20. As with the *Regen* 'rain' returns, the total number of modifiers per construction is surprisingly even. Each construction contributes approximately one-third to the overall count of modifiers (N=864). Only 22 concordance lines (4%) in total do not contain any modification (ZERO) whatsoever. Each construction also contains a similar number of COMBINATIONS of modifiers, i.e., at least two different types of modifier, NON-SPATIAL or SPATIAL, modifying the same ME.

[87] The percentages were arrived at by dividing the total per SPATIAL modifier of each construction by the total of all three constructions per SPATIAL modifier combined.

Table 59. *Distribution and aggregates frequencies of the* NON-SPATIAL *modifier variable across the three constructions (ARG+*fallen*, ARG+*geben*, and* es*+PRED) with the precipitate* Schnee *'snow'*

Construction / NON-SPATIAL modifier	*Schnee+fallen*	*Schnee+geben*	*es+schneien*	Total	Relative frequency /596
TEMPORAL	130	111	133	**374**	**63%**
QUANTITY	70	99	30	**199**	**33%**
QUALITY	38	10	43	**91**	**15%**
STANCE	35	74	91	**200**	**33.5%**
Total	**273**	**294**	**297**	**864**	**144.5%**
ZERO	3	11	8	22	4%
COMBINATIONS	146	128	129	403	68%
SPATIAL	119	70	82	271	45%

Pearson's chi-squared test: 82.096, df = 6, p-value = 1.318e-15

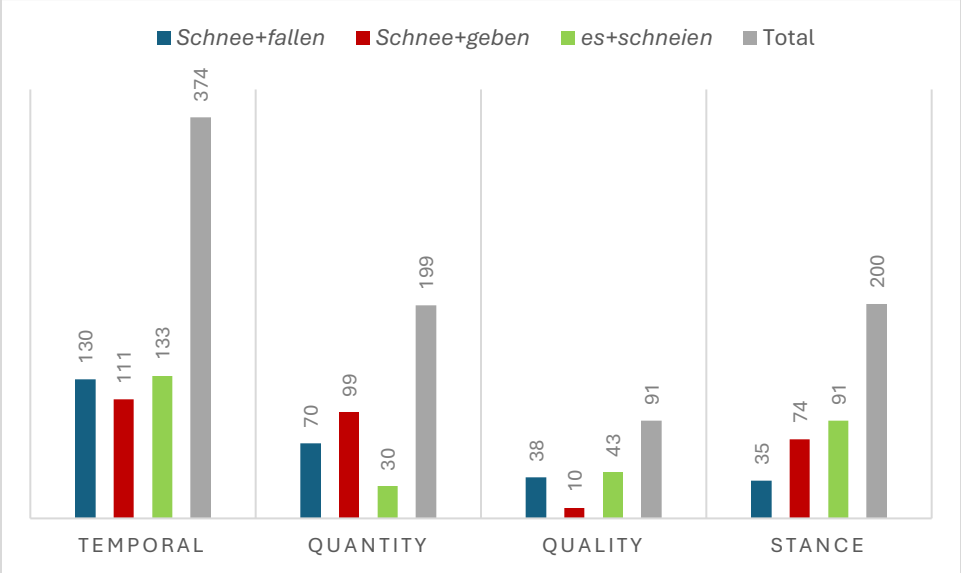

Figure 20. *Distribution and absolute frequencies of the* NON-SPATIAL *modifier variable across the three constructions (ARG+*fallen*, ARG+*geben*, and* es*+PRED) with the precipitate* Schnee *'snow'*

TEMPORAL modifiers are the most frequent per construction and overall: about two-thirds of the corpus returns for the *Schnee* 'snow' data sets contain at least one TEMPORAL modifier. Even by construction, the frequencies of the TEMPORAL modifier are almost even for the [*es+schneien*] construction and the [*Schnee+fallen*] construction (N=130) with the [*Schnee+geben*] construction being slightly less frequent. STANCE and QUANTITY are modifications that are virtually evenly distributed in the corpus returns, i.e., one-third of the data contains at least one modifier of STANCE and at least one modifier of QUANTITY. The individual differences between the constructions are striking though and appear almost inverted for [*Schnee+fallen*] and the [*es+schneien*] construction. The latter contains the highest incidence of STANCE (N=91) and

lowest incidence of QUANTITY modification (N=30) while [*Schnee+fallen*] contains the lowest amount of STANCE (N=35) and a fairly high amount of QUANTITY expressions (N=70). This dovetails the genre type findings in that the presumably more formal [*Schnee+fallen*] construction is attracted by genre types that report on quantities without an affective component. In contrast, the high incidence of STANCE associates with the overall more highly perceived informality of the [*es+schneien*] construction. In comparison, [*Schnee+geben*] contains the highest incidence of QUANTITY expressions overall (N=99) as well as a high number of STANCE modifiers (N=74), which, moreover, occur together fairly frequently (N=35). In turn, several of these instances contain negation (N=15) where the speaker expresses their positive or negative feelings about the absence of snow. The least frequent modifier overall is that of QUALITY: only 15% of the returned concordance lines (and 10.5% of all NON-SPATIAL modifiers) mark QUALITY where the tokens are chiefly shared between the [*es+schneien*] construction (N=43) and the [*Schnee+fallen*] construction (N=38). The [*Schnee+geben*] construction contributes only 10 tokens in total to modifiers of QUALITY. Having said that, the absence of quality may correlate with the prevalence of QUANTITY modifiers since these can present similar.

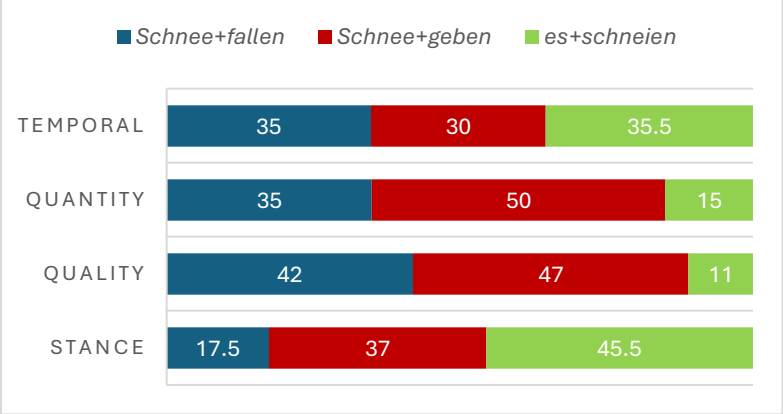

Figure 21. *Proportional contribution in percentages of construction per level of NON-SPATIAL modifier within the* Schnee *'snow' precipitate data sets*

In summary, the majority of *Schnee*-MEs contain an element of TEMPORALITY. STANCE and QUANTITY are similarly frequently expressed while QUALITY and/or MANNER only play a minor part in modification of snow events. Even so, the type of NON-SPATIAL modification, by and large, depends on the construction type, which attracts different types of modifiers to varying degrees.

3.2.5 Summary of Schnee *'snow' constructions*

The preceding quantitative and qualitative analysis of MEs employing the three constructions [*Schnee+fallen*], [*Schnee+geben*], and [*es+schneien*] in combination with precipitate *Schnee* 'snow' has provided a comprehensive account of the constructional preferences of each of the three constructions and their interaction with a number of variables. While there is great overlap between the three constructions, each construction was also found to attract different variables to varying degrees.

I carried out a Pearson's chi-squared test for Table 56 (GENRE variable), Table 57 (TAM variable), Table 58 (SPATIAL), and Table 59 (NON- SPATIAL), which revealed statistically highly significant results for each table. The chi-squared statistic (X-squared), the degrees of freedom (df), and the significance level (p-value) are listed in the following:

Table 56: X-squared = 64.816, df = 6, p-value = 4.703e-12

Table 57: X-squared = 90.193, df = 10, p-value = 4.903e-15

Table 58: X-squared = 54.453, df = 4, p-value = 4.23e-11

Table 59: X-squared = 82.096, df = 6, p-value = 1.318e-15

For Table 58, the peripheral spatial modifiers (goals, sources, and paths) were combined for each construction since otherwise the frequencies are too low for a reliable application of the chi-squared test for a significant distribution. The more general observations are summarized in the following:

The [*Schnee+fallen*] construction is used purely literally. While it also mainly attracts the WEATHER TALK genre, it contributes the lowest proportion of WEATHER TALK overall. At the same time, it is the most strongly attracted by other types of genre, especially by METEOROLOGY and by WEATHER REPORTS, which, together, account for roughly one-third of the [*Schnee+fallen*] returns. Even utterances from the LITERARY genre mostly attract the [*Schnee+fallen*] construction. The greater variety of types of genre is also mirrored in the different levels of TAM marking attracted by the [*Schnee+fallen*] construction. While MEs are expressed almost evenly in a tripartite split between PAST, EXISTENTIAL/GENERIC, and present_FUTURE, the amalgamate of "future intent" (present_FUTURE, FUTURE, IRREALIS) occurs most frequently within the [*Schnee+fallen*] construction

as well as in all of the precipitate *Schnee* 'snow' data sets, which makes the [*Schnee+fallen*] construction especially future oriented and gives it a great potentiality force. Regarding SPATIAL modification, the [*Schnee+fallen*] construction encompasses a relatively low incidence of ZERO modification and, correspondingly, not just a large number of other SPATIAL modifiers (particularly LOCATIONS) but especially peripheral SPATIAL modifiers like GOALS, SOURCES, and PATHS. These are almost solely provided by the [*Schnee+fallen*] construction within the *Schnee* 'snow' data sets, a tendency which was instantiated in the [*Regen+fallen*] returns as well. Regarding NON-SPATIAL modifiers, the [*Schnee+fallen*] construction attracts mostly TEMPORAL modifiers (and to a very similar degree as the [*es+schneien*] construction), but only a low incidence of STANCE modification, which is, in fact, the least likely to occur with the [*Schnee+fallen*] construction. This is similarly true for modifiers of QUALITY, which tend to focus on 'the first snow' (*der erste Schnee*') where they occur. Lastly, both, modifiers of TIME and PLACE exhibit great diversity with respect to semantic categories, types, and tokens.

The [*Schnee+geben*] construction, too, is exclusively used in a literal manner. It associates with genre type WEATHER TALK to the same degree as the [*es+schneien*] construction, but contributes a higher number of utterances from the METEOROLOGICAL genre. The topics of the WEATHER TALK genre are furthermore very diverse and hail from a range of URLs and blogs dedicated to specific interests. The [*Schnee+geben*] construction is, furthermore, the construction with the highest incidence of concordance lines in PAST tense, followed – with a quite a margin– by utterances in PRESENT tense and present_FUTURE equally. However, all future intent tokens combined make the [*Schnee+geben*] construction the second most frequent level of TAM within the [*Schnee+geben*] returns (after PAST tense). The proclivity for expressing future intent is shared between the two ARG-Cxx. SPATIAL mentions are plainly divided between ZERO modification and LOCATIONS with the former being slightly less frequent. Having said that, the number of LOCATIONS in the wider preceding or following context, which do usually not modify the ME directly, is noticeable; so is the absence of any peripheral SPATIAL modifiers, i.e., GOALS, SOURCES, and PATHS. With respect to NON-SPATIAL modifiers, TEMPORAL modifiers are the most frequent with a focus on the semantic category of "deictics" but also "seasons", particularly *Winter*, which is not surprising considering that snow typically occurs in winter. As opposed to the other two constructions, [*Schnee+geben*] contains a large number of QUANTIFIERS, which, in turn, encompass a large number of negations (i.e., ZERO quantity). This tendency has also been revealed for modifiers of QUANTITY within the [*Regen+geben*] construction. Lastly, modifiers of

STANCE, as opposed to the [*es*+*schneien*] construction, are very frequent, mainly positive and expressions of STANCE interact with negation in such a manner that speakers turn out to bemoan the absence of snow. To finish, modifiers of QUALITY are remarkably infrequent although the paucity of QUALITY expressions may be explained by the prevalence of QUANTITY modifiers.

The [*es*+*schneien*] construction is the only construction that includes a handful of non-literal usage events. It predominantly attracts genre type WEATHER TALK. Other types of genre only occur to a much lower degree, with the LITERARY genre being the most likely to occur. The MEs are roughly equally expressed in either PRESENT or PAST tense with the latter being slightly more frequent. Both tense levels contain lots of variety regarding TAM marking, such as quasi-auxiliaries, progressive aspect, and *Perfekt* tense. At the same time, future events are the least likely to be attracted by the [*es*+*schneien*] construction. With respect to SPATIAL modification, the data are somewhat evenly split between ZERO paths and LOCATIONS in addition to only four peripheral SPATIAL modifiers. Overall, the SPATIAL modifiers can be grouped into only four semantic categories. NON-SPATIAL modifiers are predominantly TEMPORAL and, semantically, mostly refer to frequency (*wieder* 'again') and *deixis* (*heute* 'today'). Furthermore, there is frequent expression of negative STANCE. As opposed to the ARG-Cxx with precipitate *Schnee* 'snow', neither modifiers of QUALITY nor QUANTITY occur frequently. Where QUALITY is mentioned, it is usually in reference to intensity or severity of the ME.

3.3 Constructions with the precipitate *Schneeregen* 'sleet' (lit. 'snow rain')

This section examines the *es*+PRED-Cx as well as the two ARG-Cxx (ARG+*fallen* and ARG+*geben*) with the *Schneeregen* 'sleet' precipitate. What constitutes *Schneeregen* 'sleet', appears to differ depending on the country/region it occurs in (e.g., rain that freezes as it falls from the sky or wet, partly melted falling snow).[88] As per the German definition, *Schneeregen,* literally 'snowrain', is a mixed precipitate of snow and rain where the mixing ratio may differ and either snow or rain may prevail or where snow and rain fall at the same time. As already alluded to earlier, the fact that German has a proclivity for compounds means that additional modification may be pre-empted, but it also means that these precipitation compounds are

[88] Source: dictionary.cambridge.org/de/worterbuch/englisch/sleet (last accessed December 10, 2023)

more likely to appear in ARG-Cxx than in *es*+PRED-Cxx, which is certainly also true for *Schneeregen* 'sleet'. Where the ARG is responsible for the semantic encoding of the precipitate, the predicate will be light, so that *Schneeregen* 'sleet' will presumably occur in either of the ARG-Cxx rather than the *es*+PRED-Cx. Having said that, it is conceivable for *Schneeregen* 'sleet' to occur in an ARG-PRED-Cx, such as *es regnet/schneit Schneeregen* 'it's raining/snowing sleet', where the semantic heavy lifting is split between the argument and the predicate (cf. Chapter 4, Section 4.2.2).

3.3.1 The argument construction [Schneeregen+fallen] *'sleet+fall'*

This subsection concerns the ARG-Cx [*Schneeregen+fallen*] 'sleet+fall'. The queries in the *deTenTen13* were carried out using the expressions in (xix) and (xx) to account for the different word orders. The search for the below queries returned 312 concordance lines in total. Query (xix) is roughly one-third more frequent than (xx) with 180 over 132 tokens.

(xix) [lemma="fallen"] [ ] {0,8} [word="Schneeregen"] within <s/> (N=180)
(xx) [word="Schneeregen"] [ ] {0,8} [lemma="fallen"] within <s/> (N=132)

Twelve concordance lines overall were removed as false returns.

3.3.1.1 USAGE (LITERAL VS FIGURATIVE) AND GENRE VARIABLES

With respect to variable USAGE, all concordance lines reflect expressions being used literally. There is no figurative use, nor is there negation of any kind. The distribution and frequencies for the GENRE variable as annotated for the [*Schneeregen+fallen*] construction are given in Table 60. As opposed to the corpus returns surveyed up to this point, *Schneeregen* 'sleet' is the first precipitate that is not chiefly attracted to genre type WEATHER TALK. In fact, more than half of the data (N=105; 52.5%) are, interestingly, WEATHER REPORTS, while WEATHER TALK is only the second most frequent genre type (N=83; 41.5%). Obviously, weather reports often give a list of possible precipitation types, so that *Schneeregen* 'sleet', in utterances coded as WEATHER REPORT, always appears with some byproduct or co-ME, while instances of WEATHER TALK appear either with (N=49) or without (N=35) a co-ME. Moreover, there are a few utterances from the LITERARY genre (N=10; 5%). The lowest incidence (N=2; 1%) are concordance lines referring to genre type METEOROLOGY.

Table 60. *Raw and relative frequency of the* GENRE *variable in the ARG+*fallen *construction with the precipitate* Schneeregen *'sleet' (abbreviated as SR)*

Word order/ GENRE	PREC1 *SR+fallen*	PREC2 *fallen+SR*	Total	Relative frequency /200
WEATHER TALK	51	32	**83**	**41.5%**
WEATHER REPORT	41	64	**105**	**52.5%**
METEOROLOGY	2	0	**2**	**1%**
LITERARY	6	4	**10**	**5%**
Total	**100**	**100**	**200**	**100%**

The PREC2 word order is associated exactly twice as many times with genre type WEATHER REPORT than with WEATHER TALK. This is due to the fact that many WEATHER REPORT sentences begin with a modifier of time or place, which moves the argument after the verb, which, again, has to remain in second position. The PREC1 order is associated more with WEATHER TALK than with WEATHER REPORTS by 10 concordance lines. Furthermore, the PREC1 order contains each two concordance lines more of the overall infrequent LITERARY and METEOROLOGY genre types than the PREC2 order.

3.3.1.2 The TAM (TENSE, ASPECT, MOOD) variable

Regarding the frequencies of the TAM modifier variable, as given in Table 61, the *Schneeregen* 'sleet' constructions represent another departure from the majority of constructions and precipitates in that PAST is not the most frequent tense. The most frequent tense is present_FUTURE (N=90; 45%) where the predicate is inflected for present tense, but the temporal modifiers make clear that some future event is predicted. In approximately one-third of these concordance lines (N=24), the predicate is accompanied by a modal verb, mostly *kann* 'can' (N=16). The "future intent" tokens combined make up almost half of the *Schneeregen+fallen* corpus returns (N=97; 48.5%). The second most frequent tense is PRESENT (N=53; 26.5%). The concordance lines are almost evenly distributed between the two word orders. Slightly less than half of the PRESENT tense concordance lines additionally have progressive force, as in (117), where the adverb *grade* (colloq., *gerade* 'right now'), in the second clause, indicates that the first clause is happening concomitantly.

Table 61. *Raw and relative frequency of the* TAM *variable in the ARG+*fallen *construction with the preciptate* Schneeregen *'sleet'*

Word order/ TAM	PREC1 *SR+fallen*	PREC2 *fallen+SR*	Total	Relative frequency /200
PAST	24	20	**44**	**22%**
PRESENT	26	27	**53**	**26.5%**
present_ FUTURE	26/12	50/2	**90**	**45%**
FUTURE	7	0	**7**	**3.5%**
IRREALIS	0	0	**0**	**0%**
EXIS_GEN	5	1	**6**	**3%**
Total	100	100	**200**	**100%**

Take note that example (117) additionally contains the pronoun *es* 'it' (*es fällt Schneeregen* 'it falls sleet'), which is obligatory here, unless the locative *in Berlin* is moved to the front of the clause (*in Berlin fällt Schneeregen* 'in Berlin sleet is falling') or the precipitate is moved to the front of the clause ('*leichter Schneeregen fällt in Berlin* 'light sleet is falling in Berlin'), which is by no means ungrammatical, but does sound more literary and does not necessarily "match the style" of the utterance.

(117) *Ach übrigens,* ***es fällt leichter Schneeregen*** *in Berlin, ich wäre auch lieber grade irgendwo, wo es warm ist, auch wenn da nur halbtote Backpacker sind.* (dragstripgirl.de)
'Oh, by the way, **it is falling light sleet** in Berlin, I, too, would rather be somewhere where it is warm right now, even if there are only half-dead backpackers.'

Generally, for the concordance lines annotated as progressive, it is always clear that the weather event is ongoing. For the concordance lines annotated as PRESENT, it is possible that some of these indicate future events as well. However, it was not possible to determine from the immediate context whether the speaker indicates a future event or a current event that is going to last. Utterances in the PAST tense are the third-most frequent (N=44; 22%) and are fairly evenly distributed between the two different word orders. There is no variation within the PAST tense data, i.e., no *Perfekt* or past perfect, no progessives; neither are there any quasi-auxiliaries. EXISTENTIAL/GENERIC tokens have the lowest incidence (N=6)The lack of FUTURE tokens in the PREC2 order is due to word order restrictions.

With respect to the interaction of the TENSE variable with GENRE, it turns out that the concordance lines coded as present_FUTURE predominantly belong to the WEATHER REPORT genre (N=82), which makes sense considering that WEATHER REPORTS deal with predictions. Only eight tokens are coded as WEATHER TALK. It should be noted that 19 of the WEATHER REPORT concordance lines come from two sources only (12 tokens from *wettermeldung.de* 'weather report', and

seven tokens from *vol.at* 'voralberg online news and weather') and contain very similar, boilerplate wording

3.3.1.3 The SPATIAL modifier variable

The distribution and frequencies for the SPATIAL modifier variable as annotated for the [*Schneeregen+fallen*] construction are given in Table 62. More than half of the corpus returns contain a LOCATION (N=110; 55%). Out of these, only 10 tokens (5%) do not directly modify the precipitation event. One quarter of the returns do not contain any SPATIAL modification whatsoever (ZERO), which is a relatively low number compared to the *Regen* 'rain' returns, where all three concordances contained a considerable amount of ZERO SPATIAL modification. A substantial number of utterances also contain a GOAL (N=25; 12.5%) and or a SOURCE (N=20; 10%). Together with PATHS, which have the lowest incidence (N=3; 1.5%), these more peripheral SPATIAL modifiers are almost as frequent as the utterances without any modification. These SPATIAL modifiers often occur combined: seven LOCATIONS are combined with either a SOURCE (N=2) or a GOAL (N=5). An eighth combination is made up by a PATH plus a GOAL.

Table 62. *Raw and relative frequency of the SPATIAL modifier variable in the ARG+*fallen *construction and the preciptate* Schneeregen *'sleet'*

Word order/ SPATIAL modifier	**PREC1** *SR+fallen*	**PREC2** *fallen+SR*	**Total**	**Relative frequency** /200
LOCATION	54	56	**110**	**55%**
GOAL	15	10	**25**	**12.5%**
SOURCE	15	5	**20**	**10%**
PATH	3	0	**3**	**1.5%**
ZERO	17	33	**50**	**25%**
Total	**104**	**104**	**208**	**104%**
COMBINATIONS	3	5	8	19

There are 37 stand-alone modifying phrases indicating SOURCES and GOALS in addition to the already mentioned combinations. The most frequently included GOAL is *auf die Erde* 'onto the earth' (N=3). Similar to this pattern GOALS are also given as *auf die Straße* 'onto the street', *auf den Rosa-Luxemburg-Platz* 'on Rosa-Luxemburg-Square', *auf den kleinen Ort Klingenberg in Unterfranken* 'on the little town of Klingenberg in Lower Franconia', *auf Blumen, die an einer Wand lehnen* 'onto flowers that lean against a wall' and lastly the combination of GOAL and PATH: *auf die Stadt hinab* 'down onto the city'. A few GOALS contain altitude in meters, as in (118).

(118) *Auf den Bergen wieder kälter. Die NULL-GRAD-GRENZE liegt zwischen 1000 und 800 Meter. Schnee bzw.* ***Schneeregen fällt bis 800 und 600 Meter****. Auf den Bergen lebhafter Westwind.* (vol.at 'voralberg-online')
'Colder again in the mountains. The freezing point is between 1000 and 800 meters. Snow or **sleet falls to 800 and 600 meters**. Brisk westerly wind on the mountains.'

It should be noted that several of the PREC1 order GOALS (N=7) come from the same source (*vol.at*, which is a news and weather portal for the alpine *Voralberg* region in Austria, its westernmost state). More specifically, all GOALS that include a numeral/elevation (e.g., *bis 2000 and 1800 Meter* 'down to 2000 and 1800 meters') can be ascribed to *vol.at* weather forecasts, which furthermore appear to be rather formulaic.[89] Since *Voralberg* is a popular winter sports destination, we can safely assume that the elevation and corresponding precipitation is of importance for skiers, snowboarders, and climbers. Moreover, what is characteristic about these concordance lines, is their nominal or telegram style; the very first and last sentence do not contain any verb whatsoever. This nominal style is quite typical for weather reports (cf. Liégeois 2021). In fact, another predicate-less clause like *Schneefall bis 800 Meter* 'snowfall down to 800m' would not be surprising at all in this context. The nominal style has possibly been dispensed with here due to what would have been a much less natural and slightly odd *Schneeregenfall* 'sleet fall'. Again, while one item is part of the *Duden* inventory (*Schneefall* 'snow fall'), the other is not (**Schneeregenfall* 'sleet fall'). As usual, we find exceptions to the rule, though, and the *deTenTen13* returns nine results containing the items *Schneeregenfall* (SG) *or Schneeregenfälle* (PL).

Other GOALS similar to the ones sourced from the *vol.at* URL (cf. fn.89), are *bis in tiefe Lagen* 'to low altitudes' (N=3) and, including a PATH (*bis in tiefere Lagen hinein* 'down to lower altitudes') as well as the slightly differently worded *bis in die Niederungen* 'to the lowlands'. These concordance lines do not just have a similar wording in common, but also that they all belong to the WEATHER REPORT genre and are almost exclusively expressed in present_FUTURE. All of the GOALS mentioned at the start of this section (PPs starting with *auf* 'on/onto' as well as one starting with *ins* 'into') are coded as WEATHER TALK, except for two where one is about HWW, i.e., the METEOROLOGY genre, and the other is from a religious, i.e., LITERARY source (*auf die Erde* 'onto

[89] The seven GOALS given in the vol.at concordance lines are *bis 2000 und 1800 Meter*, *bis 800 und 600 Meter*, *bis ~1000 Meter, bis ~800 Meter*, *bis ~700 Meter, bis 2600 Meter*, *bis gegen 2400 und 2200 Meter inneralpin* ,down to… inner Alpine'.

the earth'). With respect to the LOCATIVES, these are, in line with the GOALS, often indicators of "altitude", such as *im Tiefland* 'in the lowlands' or *in tiefsten/höheren* Lagen 'in the lowest/higher altitudes', but also "cardinal directions", such as *im Süden und Osten* 'in the south and the east'. "Deictis" are not that frequent but contain the most frequent type of the returns: *draußen* 'outside' (N=8). That means, that the SPATIAL modifiers here more closely resemble Liégeois's (2021) findings for weather reports, which dovetails the finding that genre type WEATHER REPORTS has the highest incidence in the *Schneeregen+fallen* returns.

3.3.1.4 The NON-SPATIAL modifier variable

The frequencies and distribution of NON-SPATIAL modifiers across the [*Schneeregen+fallen*] construction are given in Table 63. SPATIAL modifiers are listed in the last row and count towards ZERO modification and combinations. That means, an utterance with a SPATIAL modifier and a NON-SPATIAL modifier counts as a combination and a concordance line with a SPATIAL modifier only does not count as ZERO even if it does not contain a NON-SPATIAL modifier. The most frequent type of NON-SPATIAL modifier is TEMPORAL (N=124; 62%). Modifiers of QUALITY or MANNER have the second highest incidence. Modifiers of QUANTITY and STANCE are similarly frequent with 30 and 26 tokens respectively (15% and 13%). In more than two-thirds of the concordance lines, two or more SPATIAL and/or NON-SPATIAL modifiers combined. There is only a small number of five concordance lines which contain no modifier whatsoever (ZERO).

***Table 63**. Raw and relative frequencies of the NON-SPATIAL modifier variable in the ARG+*fallen *construction with the preciptate* Schneeregen *'sleet'*

Word order/ NON-SPATIAL modifier	**PREC1** *SR+fallen*	**PREC2** *fallen+SR*	**Total**	**Relative frequency** /200
TEMPORAL	53	71	**124**	**62%**
QUANTITY	13	13	**26**	**13%**
QUALITY	26	15	**41**	**20.5%**
STANCE	19	11	**30**	**15%**
Total	**111**	**110**	**221**	**110.5%**
ZERO	3	2	5	2.5%
COMBINATIONS	72	65	137	68.5%
SPATIAL modifier	82	69	151	75.5%

The most frequent combination (N=42) of two modifiers only is that of NON-SPATIAL modifier (TEMPORAL) with a SPATIAL modifier, which is fairly evenly distributed between the PREC1 order (N=20) and the PREC2 order (N=22). These frequencies were counted separately and are not

listed in the table. The following subsections elaborate on each modifier individually, starting with modifiers of TIME.

Modifiers of TIME. There are 124 TEMPORAL modifiers in total, meaning that 62% of the returns appear with some sort of temporal modification. Fairly infrequently (N=17), the TEMPORAL mention occurs as the only type of modification. Where it appears in combination with at least one other modifier, it is predominantly a modifier of LOCATION (N=87, contained in combinations of at least two modifiers). Regarding their semantic makeup, the modifiers can largely be categorized as detailing "day of week", "time of day", "frequency/duration", "sequence", and "other/combinations", which is mostly a catchall for combinations of the former categories. "Time of day" is the most frequent category (N=15) and mostly these are instances of *in der Nacht* 'during the night'. This modifier is presumably frequent because outside temperature tends to drop overnight, which may lead to snow or sleet fall, rather than rain, especially regarding forecasts. Regarding PAST tense utterances, like snow, speakers will be able to notice remnants of the precipitate the next morning and report on it. Other "time of day" modifiers are such as *am Nachmittag* 'in the afternoon', *ab 14 Uhr* 'from 2pm', or *im Tagesverlauf* 'during the course of the day'. The category "day of the week", e.g., *am Montag* 'on Monday' is fairly frequent, too, (N=14); sometimes, the exact date is added (N=3). Other "day of the week" examples indicate a timepoint that starts a timespan, for example, by using preposition *ab* 'from' rather than *am* 'on'. Again, in some concordance lines, the modifier is provided in the periphery rather than directly modifying the event as, for example, the modifier *Samstag* 'Saturday' in (119). Clearly, the speaker has plans for the day in question, which the weather could negatively affect.

(119) *Und es wird immer kälter. In den Höhenlagen des Sauerlandes* ***soll Schneeregen fallen****. Das sieht für Samstag gar nicht so gut aus*. (wa.de 'Westphalian Gazette')
'And it is getting colder. **Sleet is expected to fall** in the higher elevations of the Sauerland region. It does not look too good for Saturday.'

Often, the TEMPORAL modifier represents a combination of weekday and time of the day, again, mostly specifying the night leading up to the mentioned weekday as in *die Nacht zum Donnerstag* lit. 'the night [leading up] to Thursday', i.e., 'Wednesday night'. In example (120), the decreasing temperatures are modified by *am Freitag* 'on Friday' and the precipitation event is modified by *über den ganzen Tag verteilt* 'spread out throughout the day'. However, *throughout the day* obviously refers to the *Friday* mentioned earlier. Strictly speaking, this is a

"time of day" modifier, but the connection to "day of week" also makes it a combination of temporal modifiers.

(120) *Schon am Freitag sinken die Temperaturen wieder leicht, außerdem* ***fällt über den ganzen Tag verteilt*** *etwas Regen oder* ***Schneeregen***. (blogspot.de)
'Temperatures will drop again slightly on Friday, and some rain or **sleet will fall throughout the day.**'

The most frequent deictic TEMPORAL modifier is *heute* 'today' (N=4). Other examples are *im Moment* 'at the moment', *inzwischen* 'meanwhile', or *eine Woche später* 'one week later'. The fact that there are often different weather events as one moves through a day, is covered by the smallest category of temporal modifiers: "sequence" (N=8). The sequence of weather events is specified, for example, as *anfangs* 'at the beginning', *zuerst* 'first', *nachfolgend* 'subsequently, or *letztendlich* 'finally'. The largest category "other" (N=27) includes a large number of combinations of different categories but also items of frequency (*gelegentlich* 'occasionally', *meistens* 'mostly') and rather idiosyncratic temporal modifiers such as *als ich in den Garten ging* 'when I went into the yard', *am Tag nach dem Fernsehinterview* 'the day after the TV interview, and *recht pünktlich zum Ende der Fahrerbesprechung* 'pretty much on time for the end of the drivers' briefing' but. The combinations can be simpler combinations such as *in der Nacht zum Mittwoch.... gelegentlich* 'Tuesday night… occassionally', but also more complex ones such as *am Mittwoch.... besonders am Morgen sowie am Vormittag … zeitweise* 'on Wednesday.... especially in the morning and in the morning … at times' as well as the one given in example (38) in Chapter 2, Section 2.2.4, where I added the semantic category of the temporal modifier in superscript directly after the respective modifier in the translation.

The example in (38) is the most complex in terms of its combinations of TEMPORAL modifiers, which include "sequence", "time of day", "weekday", an "exact date" and at least one "season" plus a "time period". Liégeois (2021) found that a nominal style dominates temporal modification, i.e., nouns like days of the week or time of the day are more frequent than adjectives or adverbs. This seems true for the [*Schneeregen+fallen*] construction's temporal modifiers as well, as the weekdays and time of the day (as well as their combinations) alone add up to 50 nominal style modifiers. There is only one token each of the adverbs *nachmittags* 'in the afternoon' and *abends* 'in the evening'. Otherwise, the ME's "time slots" are filled with PPs (e.g., *am Samstagnachmittag* 'on Saturday afternoon'). However, rather than having clear-cut

categories, there are many combinations as well as rather idiosyncratic points in time like the ones specified as "other".

Modifiers of QUANTITY. Modifiers of QUANTITY exhibit the lowest incidence (N=26; 13%) of all NON-SPATIAL modifier types and are evenly distributed between the two word orders. Only one of these modifiers is a stand-alone modifier (*etwas* 'a little'); the other are combined with at least one other modifier, which mostly is a SPATIAL modifier (N=19) or a TEMPORAL modifier (N=17), or a combination of both. Regarding the type of QUANTITY modifier, there are basically only two categories: "small quantities" and "large quantities", where small quantities prevail. The stand-alone modifier *etwas* 'a little' is also the most frequent type overall and accounts for more than half of the data (N=15). The adjective *gering* 'a small amount of/slight' accounts for three tokens. Furthermore, there are synonyms like *ein bisschen* 'a little' und *ein wenig* 'a little'. The modifiers denotating small amounts are all used to directly modify *Schneeregen* 'sleet' or a combination of two precipitates (*etwas/ein bisschen Regen und Schneeregen* 'some/a little rain and sleet'). The quantifiers which specify a larger amount do not directly modify *Schneeregen* 'sleet' but refer to precipitation that falls *als Schneeregen* 'as sleet', as in (121). Strictly speaking *einiges* 'plenty/quite a bit' modifies *Schneefall* 'snowfall', however, ultimately, the quantifier encompasses *Schneeregen* 'sleet' in this case.

(121) ***Es gibt einiges an Schneefall**, der in tiefen Lagen durch die gestiegenen Temperaturen **als Schneeregen** vom Himmel **fallen** kann.* (wissenswertes.at 'thingstoknow')
'**There is (a) considerable (amount of) snowfall** that, at low altitudes, can **fall** from the sky **as sleet** due to the increased temperatures.'

Furthermore, there is the unique quantifier type *mehr und mehr* 'more and more', which modifies a list of precipitates (*mehr und mehr Nassschnee, Schneeregen bzw. Regen* 'more and more wet snow, sleet, respectively rain') as well as the quantifier *enorm* 'enormous'. Except for one instance, all of the QUANTITY modifiers interact with the WEATHER REPORT genre with regard to the PREC2 order. In the PREC1 order, four QUANTITY modifiers (all *etwas* 'some') interact with genre WEATHER REPORT and five interact with genre WEATHER TALK (two tokens are *etwas* 'some'). That means, two-thirds of the quantifier data interact with genre WEATHER REPORT.

Modifiers of QUALITY. One-fifth of the NON-SPATIAL modifiers mark QUALITY (N=41; 20.5%), mostly contained in the PREC1 order (N=26). Only seven of these modifiers stand alone without any other modifier. Where QUALITY is combined with another modifier, it is either a LOCATIVE (N=32) or a TEMPORAL modifier (N=30), or both (N=18). All in all, there are numerous types and

unique tokens, which can be sorted into four or five semantic categories. The most frequent modifier type is *leicht* 'light' (N=6) and, in three cases, it is *leichter Niederschlag* 'light precipitation' that falls *als Schneeregen* 'as sleet'. Overall, there are eleven tokens where the QUALITY modifier strictly speaking relates to *Niederschlag* 'precipitation' or *Schauer* 'showers' that falls *as* sleet. Another eleven tokens focus on properties of the sleet (*matschig* 'slushy', *garstig* 'nasty', *gefrierend* 'freezing', *dicht* 'dense', *körnig* 'grainy', etc., which could be grouped with all instances of *leicht* 'light') and eight tokens refer to the manner of falling (e.g., *sehr schnell* 'very fast', *rieselnd* 'fluttering', *in Strömen* 'in streams', *sacht* 'softly', *bleiern* 'leaden'). Lastly, five tokens are "other" and specify that the quality was *teils/teilweise* 'partly' sleet, that *so was Ähnliches* wie 'something akin to' sleet fell, or that the quality is such that 'it will not stay' (*wird dieser wohl nicht liegen bleiben*).

Modifiers of STANCE. The corpus returns representing my data sample contain 30 concordance lines expressing almost exclusively a negative attitude of the *Schneeregen* 'sleet' event. Only one utterance is positive and the sleet is described as *erfrischend* 'refreshing' while the atmosphere is being juxtaposed as *angeheizt* 'heated'. Another utterance expresses uncertainty in a conditional *niemand wusste so recht, ob Schnee oder Schneeregen fallen würde* 'nobody quite knew whether snow or sleet would fall'. Generally, the expression of negativity appears hedged (cf. ex. (119) above). However, the more informal the weather talk, the less hedged the disdain tends to be (e.g., *bäh* 'bah', *das Wetter ist mies* 'the weather is nasty', *ekliger/garstiger Schneeregen fiel* 'disgusting/nasty sleet fell'). More often, the negative prosody can be found in the left and/or right context, though, as in (122), where it occurs in the right context only.

(122) *Zuerst **meinte es der Wettergott nicht so gut** mit uns und **es fiel ein Schneeregen**.*
(strohgaeunarren.de 'Strohgäu jesters')
'At first, **the weather god did not mean well** with us and **it fell a sleet**.'

Apart from this example invoking a weather god, it is also intriguing in that it contains an indefinite article (*ein* 'a') but no other modifying material, which is not possible in English (*a sleet fell), as well as *es* ('it', *it fell sleet/rain/snow). The second main clause (*und es fiel ein Schneeregen*, lit. 'and it fell a sleet') sounds slightly odd but not altogether unacceptable. The source website belongs to a carnival guild in the state of *Baden-Württemberg* (in the very south-west of Germany) and it is possible that the use of an indefinite article before a precipitate is a

regional idiosyncrasy. The next section surveys the second ARG-Cx [*Schneeregen+geben*] 'sleet+give'.

3.3.2 The argument construction [Schneeregen+geben] *'sleet+give'*

This section examines the [ARG+*geben*] construction with the precipitate *Schneeregen* 'sleet'. The search in the *deTenTen13* for the query in (xxi) returned 176 results (two concordance lines were discarded) and the query for (xxi) returned 192 results (15 concordance lines were discarded), which leaves 351 concordance lines in total. The original query included the pronoun *es* 'it' and returned almost the same frequencies as those given in Table 64 with a deviance from the now used data of four to five items only.[90] Since *es* 'it' is obligatory in the [ARG+*geben*] construction anyway, running the query without it allowed for the pronoun to show up in positions not specified in the regular expression. As expected, all returns contain *es* 'it', which does show up in different positions (cf. Table 64). As there are two arguments in this construction —*es* 'it' and *Schneeregen* 'sleet'— none of the returns were discarded as long as (xxi) returned PREC1 and (xxi) returned PREC2 results, regardless of the position of *es* 'it'.

(xxi)	[word="Schneeregen"][]{0,8}[lemma="geben"] within <s/>	(N=176)
(xxii)	[lemma="geben"][]{0,8}[word="Schneeregen"]within <s/>	(N=192)

The results show, however, that the majority of returns followed the default word orders given in the first row of Table 64 (each with 87 returns). The exploration of the different sentence position that *es* 'it' can occur in was carried out for the *Schneeregen* 'sleet' precipitate since the

[90] Original query for the PREC1 order was "[word="es"][]*[word="Schneeregen"][]*[lemma="geben"] within <s/>" and returned 148 items.

lower overall number of returns exposed the differences of the pronoun position more clearly than other precipitates.

Table 64. *Frequencies of different sentence positions of* es *'it' in concordance lines returned for queries (xxi) and (xxii)*

Word order	PREC1 *Schneeregen+geben*	Frequency	PREC2 *geben+Schneeregen*	Frequency
	es*+Schneeregen+geben*	**87**	*geben+es+Schneeregen*	**87**
	*Schneeregen+geben+***es**	**12**	*es+geben+Schneeregen*	**13**
	Schneeregen(+AUX)+***es****+geben*	**1**		
Total		**100**		**100**

There are seven negations in the returned concordance lines. The following subsections examine the different variables starting with the variables USAGE and GENRE.

3.3.2.1 USAGE (LITERAL VS FIGURATIVE) AND GENRE VARIABLES

As with the [*Schneeregen+fallen*] construction, there are no instances of figurative usage, all returned concordance lines are used literally. The distribution and frequencies of the GENRE variable within the [*Schneeregen+geben*] data are given in Table 65. The concordances lines, except for three instances, are almost exclusively split between genre types WEATHER TALK and WEATHER REPORT. The former is the most frequent with more than three quarters of the returns (N=154; 77%) while the latter contains roughly one-fifth of the returns (N=43; 21.5%). There are two concordance lines from the LITERARY genre and a single concordance line referring to METEOROLOGY.

Table 65. *Raw and relative frequencies of the GENRE variable within the ARG+geben construction and preciptate* Schneeregen *'sleet' (SR)*

Word order / GENRE	PREC1 *SR+geben*	PREC2 *geben+SR*	Total	Relative frequency
WEATHER TALK	69	85	**154**	**77%**
WEATHER REPORT	29	14	**43**	**21.5%**
METEOROLOGY	1	0	**1**	**5%**
LITERARY	1	1	**2**	**1%**
Total	**100**	**100**	**200**	**100%**

Looking at the constructions by word order, both are similar in that WEATHER TALK makes up the lion's share of concordance lines, although the PREC2 order has a higher frequency count for utterances in the past tense. The PREC1 order, on the other hand, has more than twice as many concordance lines coded as WEATHER REPORT due to word order rules. The PREC1 WEATHER REPORTS

contain more modal verbs, as we shall see in the next subsection. The LITERARY and METEOROLOGICAL genres are noteworthy in that that they are barely present with only three tokens in total.

3.3.2.2 The TAM (TENSE, ASPECT, MOOD) variable

The distribution and frequencies of the TAM variable within the *Schneeregen+geben* corpus returns are given in Table 66. As opposed the *Scheeregen+fallen* returns, the frequencies here continue the tendency we have seen with other precipitates: weather events are predominantly talked about in the PAST tense (N=105; 52.5%). The second highest incidence contains "future intent" tokens, i.e., present_ FUTURE (N=65; 32.5%) and future tense (FUTURE, N=10; 5%), which together add up to 75 tokens (37.5%). In contrast to the [*Schneeregen+fallen*] construction, MEs in PRESENT tense are fairly infrequent (N=17; 8.5%); roughly half of the tokens are additionally marked for progressive aspect. The lowest incidence is provided by concordance lines that express EXISTENTIAL or GENERIC meaning (N=3; 1.5%). Word order differences and variations within the respective level of tense are elaborated on below the table.

Table 66. *Raw and relative frequencies of the TAM variable in the ARG+*geben *construction and the precipitate* Schneeregen *'sleet' (SR)*

Word order/ TAM modifier	**PREC1** *SR + geben*	**PREC2** *geben + SR*	**Total**	**Relative frequency** /200
PAST	34	71	**105**	**52.5%**
PRESENT	6	11	**17**	**85%**
present_FUTURE	48	17	**65**	**32.5%**
FUTURE	10	0	**10**	**5%**
IRREALIS	0	0	**0**	**0%**
EXIS_GEN	2	1	**3**	**1.5%**
Total	100	100	**200**	**100%**

Looking at the concordance lines by word order, we are faced with mixed results. As with the overall trend (compare the rain and snow precipitates), PAST clearly dominates the PREC2 returns (N=71; 35%), followed by present_FUTURE (N=17; 8.5%), but there are no utterances in the periphrastic FUTURE tense and no presence of modal verbs. By contrast, PREC1, for most of the results, does not mirror the frequencies of the PREC2 order. Here, PAST only has the second highest incidence (N=34; 17%), which is less than half as many for the PREC2 order, while almost half of the data appear in present_FUTURE (N=48; 24%). These concordance lines overwhelmingly contain some type of modal verb (N=41), which explains the PREC1 word order. Mostly, these

modals are *kann* 'can' (N=24) and *könnte* ('can/could', N=2), as well as forms of *sollen* ('supposed/should to', N=15). Generally, likelihood/certainty is higher with *kann* 'can' than *könnte* 'could'. Additionally, all FUTURE tokens are contained within the PREC1 order (due to word order rules). The two word order distributions converge in having very few returns with existential or generic meaning.

Overall, there is some variation within the PAST tense data, which includes six utterances in the *Perfekt* tense and five utterances in the past perfect. Three concordances contain some type of quasi-auxiliary or at least express that the start of the rain coincided with the start of another event (N=1) or that the ME stopped at some point. Examining the FUTURE concordance lines more closely, they are furthermore hedged by particles and adverbs expressing probability/possibility such as as *wohl* 'perhaps', *wahrscheinlich* 'probably', *vielleicht* 'maybe', *eher* 'rather', and subordinating conjunction *ob* 'whether'. Correspondingly, out of the concordance lines that contain some type of future marker, 39 instances present with some type of co-ME in the same construction, usually, because, in a forecast situation, the forecaster cannot always be sure of the type of precipitation that is predicted to fall. Section 3.5.2.5 elaborates on the co-MEs, which occur with *Schneeregen* and appear more than twice as many times with some sort of future marker (N=30) than with other tenses or aspects (N=15).

3.3.2.3 The SPATIAL modifier variable

The distribution and frequencies of the SPATIAL modifier variable within the *Schneeregen+geben* data are given in Table 67. Regarding the aggregate frequencies of the SPATIAL modifier variable, more than half of the returns contain a LOCATION (N=109; 54.5%). Only a small number of locations (N=14; 7%) do not directly modify the ME in the immediate context but appear in the wider preceding or following context. The second most frequent are utterances with no SPATIAL mentions whatsoever (ZERO, N=86; 43%). There is a low incidence of "minor" SPATIAL modifiers, such as GOALS (N=3), SOURCES (N=4), and PATH (N=1), which mostly occur in the PREC1 order, which also contains the three combined SPATIAL modifiers present in the returns.

Table 67. *Raw and relative frequency of the* SPATIAL *modifier variable in the ARG+geben construction and the preciptate* Schneeregen *'sleet'*

Word order/ SPATIAL modifier	**PREC1** *SR + geben*	**PREC2** *geben + SR*	**Total**	**Relative frequency** /200
LOCATION	56	53	**109**	**54.5%**
GOAL	2	1	**3**	**1.5%**
SOURCE	4	0	**4**	**2%**
PATH	1	0	**1**	**0.5%**
ZERO	40	46	**86**	**43%**
Total	**103**	**100**	**203**	**101.5%**
Combinations*	3	0	3	1.5%

*Combinations: each one G+P, L+G, and S+L

The four SOURCES are variations of *aus der Wolkendecke* 'from the cloud cover', (lit. 'out of the cloud cover'). One time, the cloud cover is additionally described as *dickwattig* 'thick cottony soft'. The one combination of the PATH and the GOAL is *bis in alle Täler hinab* 'down into all the valleys'. And the combination of LOCATION and GOAL is deictic *hier* 'here' and the GOAL *auf unser Haupt* 'on[to] our heads', as shown in (123).

(123) ***Hier hat's*** *heute* ***Schneeregen****, Regen und Hagel* ***auf unser Haupt gegeben*** *und das nicht zu knapp!!* (bag-obsession-blog.de)
'Here, **it has given sleet**, rain and hail **on our heads** today and not too little!!'

The locative modifiers can be easily grouped into five semantic categories. The most frequent category, "location" is the largest. The locations are mostly cities (*in Berlin/Nizza*) and regions (*in Dithmarschen*, which is a district in the northern state of *Schleswig-Holstein*; *an den Küsten* 'on the coasts'). Accordingly, these are all nominal regarding their syntactic type. The second largest group is "deictics", which mostly contains adverbs, particularly, types *hier* 'here' (N=9) and *bei uns/mir* 'at us/me/here' (N=7). Other deictic modifiers are such as *da*/*dort* 'there', *überall* 'everywhere', *örtlich* 'locally', *draußen* 'outside'. Moreover, there are the smaller categories "cardinal direction" (*im Süden* 'in the south', *im Osten des Landes* 'in the east of the country') and "altitude" (*auf 1450m* 'at 1450m', *in tiefen Lagen* 'in low altitudes', *in höheren Lagen* 'in higher altitudes'. Lastly, the category "minor paths" contains the goals, sources and the path mentioned above. There are only few combinations of different locatives, such as "location" plus "cardinal direction", such as *im Bergland und nach Osten hin* 'in the mountains and to the east', or "deictic" plus "location" *hier in Berlin* 'here in Berlin'. As with the nominal spatial modifiers found in the *Schneeregen+fallen* data, the nominal spatial modifiers here more closely resemble Liégeois's (2021) findings for weather reports, in spite of the fairly frequent occurrence of adverbial deictics.

3.3.2.4 The NON-SPATIAL modifier variable (TEMPORAL, QUANTITY, QUALITY, STANCE)

The distribution and frequencies of various NON-SPATIAL modifiers across the ARG+*geben* construction with precipitate *Schneeregen* 'sleet' are given in Table 68. There are only four concordance lines without any modifiers and 278 NON-SPATIAL modifiers in total. SPATIAL modifiers are listed in the last row and contribute to rows ZERO and COMBINATIONS. Roughly four-fifths of the corpus returns for *Schneeregen+geben* contain a TEMPORAL modifier. The second most frequent type of NON-SPATIAL modifier expresses STANCE (N=70; 35%). Roughly half as many modifiers as STANCE mark QUALITY (N=34; 17%). As with *Schneeregen+fallen*, QUANTITY modifiers have the lowest incidence (N=16; 8%). Lastly, there is a combination of at least two modifiers, SPATIAL or NON-SPATIAL, in roughly 70% of the data.

Table 68. *Raw and relative frequencies of the NON-SPATIAL modifier variable across the ARG+*geben *construction with the preciptate* Schneeregen *'sleet' (SR)*

Word order/ NON-SPATIAL modifier	**PREC1** *SR + geben*	**PREC2** *geben + SR*	**Total**	**Relative frequency** /200
TEMPORAL	76	82	**158**	**79%**
QUANTITY	9	7	**16**	**8%**
QUALITY	16	18	**34**	**17%**
STANCE	30	40	**70**	**35%**
Total	**131**	**147**	**278**	**139%**
ZERO	2	2	4	2%
COMBINATIONS	68	71	139	69.5%
SPATIAL modifier	58	38	96	48%

The modifiers are, for the most part, somewhat homogeneously distributed between the two word orders. The following sub-sections elaborate on each modifier in turn starting with the most frequent, modifiers of TIME.

Modifiers of TIME. TEMPORAL modifiers are ubiquitous in this constructional dataset. Not only are there many tokens, but also eight semantic categories altogether: "deictics", "frequency", "time of day", "day of week", "exact date", "idiosyncratic", "combinations", and "other". "Deictics" represents the largest category (N=43) and the most frequent type therein is *heute* 'today' (14 stand-alone tokens). Other common items are *morgen* 'tomorrow', *gestern* 'yesterday', and *im Moment* 'at the moment'. The second largest category is "combinations": often these are deictics or adverbials of frequency combined with the "time of day" or "day of

week" (e.g., *heute ab Nachmittag* 'today starting in the afternoon', *hin und wieder am vergangenen Sonntag* 'on and off last Sunday'), but there are also expressions of sequentiality (*zuerst* 'first') and exact dates. Evenly distributed categories (N=17 each) are "time of day", "day of the week", and "idiosyncratic". "Day of the week" items are invariably nominal regarding PoS (e.g., *am Mittwoch* 'on Wednesday'). "Time of day" items are predominantly nominal (*um 15 Uhr* 'at 3pm', *im Tagesverlauf* 'in the course of the day') but also encompass some adverbials (*nachts* 'at night'). "Idiosyncratic items" (*auf der Rückfahrt* 'on the way back', *wenn das Wetter umschlägt und die Temperatur fällt* 'when the weather changes and the temperature drops') can occur in combinations, too, such as the following, which is preceded by a marker of duration: *kurz in der ersten Halbzeit* ('briefly in the first halftime'). Lastly, there are nine tokens which contain "exact dates" (*am 1.und 12. März* 'on March 1 and 12') and the remainder are "other" items (*heute* 'nowadays', *im Mai* 'in May'), which did not occur with enough tokens to form categories of their own.

Modifiers of QUANTITY. QUANTITY modifiers form the smallest NON-SPATIAL modifier group (N=16), which can be broken into three semantic categories: "small quantity", "large quantity", and "zero quantity (negation)". As with the [*Schneeregen+fallen*] construction, smaller quantities are the most frequent (N=7) and are most often denoted by type *etwas* 'some' (N=5). The category "large quantity" is referred to by four expressions, e.g., *viel* 'a lot' and *einigen* 'quite a bit of'. None of the QUANTITY modifiers are nominal. Seven concordance lines include some type of negation, such as the more default *kein* 'no' (N=3), *statt* 'instead of' (N=2), and *weder… noch* 'neither…nor' (n=1). One utterance states that the sleet, that otherwise fell in April, has been replaced with hail showers (*der Schneeregen, der sonst im April fiel, (ist) durch … Hagelschauer ersetzt worden*). While "xsmall quantity" and "zero quantity" are equally frequent, seven tokens is a comparatively low number of negations, particularly for the ARG+*geben* construction, which, with the precipitates *Regen* 'rain' and *Schnee* 'snow', has revealed itself to frequently encompass expressions of negation. However, since *Schneeregen* 'sleet' is a less common precipitate than *Regen* 'rain' and *Schnee* 'snow', it is less likely for people to hope that sleet, rather than rain or snow, will not fall, or, in the case of precipitation that did *not* happen to assume that it would have taken the aggregate form of *Schneeregen* 'sleet'.

Modifiers of QUALITY. The 34 QUALITY modifiers can be grouped into four categories. The largest and most easily identifiable group is that of "intensity", where the majority focus on the severity (e.g., *heftig/kräftig/stark* 'severe/hefty' N=8; *in Sturzbächen* 'in torrents' and *mit*

Eisbällchen 'with little balls of ice', each N=1) rather than the lightness (e.g., *leicht* 'light', N=4). This is contrary to the [*Schneeregen+fallen*] construction, where the focus was on the lightness. There another four tokens which specify that it is 'the first' (*den ersten*) sleet that fell, and the remainder are "other", i.e., unique types, such as *anhaltend* 'on-going', *dünn* 'thin', or *den angekündigten* 'the announced', *gefrorenen* 'frozen', *der nicht liegen blieb* 'which did not stay'. With respect to PoS, except for two nominals, the modifiers are adverbs and adjectives.

Modifiers of STANCE. Expressions of STANCE are contained in more than one-third of the corpus returns for the [*Schneeregen+geben*] construction and can be grouped into four categories. Predominantly, stance finds expression in "negative" form, such as *ich hatte keine Lust mich bei diesem Wetter im Freien aufzuhalten* 'I did not feel like being outside in this weather'. There are 44 utterances laced with negative semantic prosody, such as speakers forecasting accidents, referring to events that were spoiled by the weather, referring to the weather as *scheußlich* 'nasty', and also interjections such as *brrrr* and *bäh* 'bah/ew' (N=3) or the use of emoticons like ":(" (N=2). Five tokens are "positive", but three of these also contain negation. In one, the speaker delights in cuddling up on their couch inside while the weather outside is bad and another one is actually positive where the speaker deems sleet preferable to no snow whatsoever. "Miratives" (N=11) contain the most frequent type, *sogar* 'even' (N=10) and one token of *überraschenderweise* 'surprisingly'. The remainder are coded as "other", which encompass items where the speaker's message is not quite clear without any further context, where the speaker is indifferent (**achselzucken** 'shrug'), or where the speaker considers the ME as "not that bad", specifically marked by adverb *nur* ('only', N=4). These can either indicate disappointment that no "real snow" has fallen yet or relief that it is just sleet, which probably will not stay.

3.3.2.5 The co-ME (co-occurring meteorological event) variable

In a total of 67 concordance lines, *Schneeregen* 'sleet' occurs with at least one co-ME, which is predominantly *Schnee* 'snow' or *Schneefall* 'snow fall' (N=36). The second most frequent co-ME is *Regen* 'rain' (N=23) and eight utterances state both *Schnee* 'snow' and *Regen* 'rain' as co-MEs. Another 19 utterances contain other co-MEs, such as wind, cold, low temperatures, graupel, or they specify the transition of precipitates. The two major conjunctions joining the co-MEs are *und* 'and' and *oder* 'or'. Concordance lines containing co-MEs seem to interact with the variables TAM and GENRE. Table 69 shows the conjunctions and TAM levels used in utterances with co-ME.

The row "FUTURE / IRREALIS marker" condenses all "future intent" tokens. The second row pools any other tenses. The two major conjunctions appear with 25 tokens (*und* 'and') and 19 tokens (*oder* 'or') respectively. Other concordance lines with each one to three tokens of a conjunction are subsumed under "other" (N=10). The most frequent is the abbreviation *bzw*., i.e., *beziehungsweise* 'alternatively' (N=3).

The fact that co-MEs occur more than twice as often with utterances that contain some sort of future/irrealis marker can be explained by the fact that we do not know what will happen in the future. The second clause in example (124) below gives *Regen* as a co-ME of *Schneeregen*, while the first clause additionally contains a different argument construction ([*Schnee+fallen*]). The [*Schnee+fallen*] construction names a different location and a different precipitate than the second construction with *geben* 'give', which identifies two locations and two possible precipitates. The second construction, which is the main clause, furthermore, includes modal verb *könnte* 'could' as opposed to the subordinate *während*-clause ('while'), which indicates that the forecast for the state of Bavaria appears to be predictable with some certainty.

(124) *Während in Bayern zeitweise* ***Schnee fällt, könnte es*** *Richtung Rhein und Schwarzwald nochmals* ***Schneeregen und Regen geben*** (wettertopia.de – 'weathertopia')
'While, occasionally, **snow falls** in Bavaria, **it could give sleet and rain** again towards the Rhine and Blackforest.'

Table 69. *Frequency of conjunctions for nominal co-MEs of the precipitate Schneeregen 'sleet' by TAM factor*

Connector/ TAM	***oder*** 'or'	***und*** 'and'	**other***	**Total**
with FUTURE / IRREALIS marker	17	15	7	**39**
other	2	10	3	**15**
Total	**19**	**25**	**10**	**54**

**bzw*. 'respectively' (N=3), and each one token of: *eher… wobei* 'rather…but', *mit* 'with', *weder… noch* 'neither…nor', *neben… auch* 'besides…also', *mit* 'with', *zuerst…dann* 'first…then', slash (/)

Obviously, when speakers recount events which happened in the past or that take place in the moment of speaking/writing, they know exactly the type of precipitation they experienced or are experiencing (cf. ex. (123), p. 220). This may be more than one type of weather event at the same time, or a change of outside temperature may turn rain into snow or vice versa with *Schneeregen* 'sleet' as a continuous or intermediate aggregate phase. There may also be different types of weather in the same region due to different altitudes. Obviously, it will be

colder higher up and warmer lower down, so that the same precipitation may come down as snow higher up on a mountain and come down as rain further down on the mountain. Taking these shifts and geographical differences into account, weather forecasters may not be able to predict the type of precipitate reliably and therefore list at least two types, and possibly several, as in example (124). Correspondingly, almost all of the WEATHER REPORT tokens, except for one, occur with a co-ME.

3.3.3 The predicate construction [*es+schneeregnet*] *'it sleet' (lit. 'it snowrain')*

This section deals with the *es*+PRED-Cx construction and the precipitate *Schneeregen* 'sleet'. The queries in (xxiii) and (xxiv) were carried out to obtain the data coded and analysed in this subset. The *deTenTen13* queries for the predicate construction did not yield sufficient concordance lines (N=88 instead of N=200), which was expected. The verb *schneeregnen* 'sleet' has not yet been entered into the *Duden*, which reflects that it has not yet managed to become established in everyday language use. That means, it does not occur frequently enough to warrant an entry in the *Duden*. In this light, 88 returns in total is still more than anticipated. Furthermore, out of the 88 concordance lines, there was not a single one that had to be discarded due to being a false return or that was being used figuratively.

(xxiii) [word="es"] [ ] {0,8} [lemma="schneeregnen"] within <s/> (N=53)
(xxiv) [lemma="schneeregnen"] [ ] {0,8} [word="es"] within <s/> (N=35)

The predicate construction [*es+schneeregnet*] 'it's+sleeting' is marked in that the verb *schneeregnen* has not yet been added to the *Duden* and can therefore be regarded as emergent and highly informal. While some speakers outright reject the use of the verb *schneeregnen*, there are others who readily accept it, and some who spearhead its usage. This assessment is absed on a previous study (Heinrich 2017), in which I ran an anonymised acceptability questionnaire and furthermore asked native speakers about their intuitions regarding the use of *schneeregnen* 'to sleet' as a verb. As there were almost one hundred concordance lines altogether, they were downloaded into one data file annotated in the established manner. The distribution between word orders has not been taken into account due to the overall low number of returns.

3.3.3.1 *USAGE (LITERAL VS FIGURATIVE) AND GENRE VARIABLES*

In Table 70, the frequencies of the two variables GENRE and TAM are displayed. As already noted, there are no instances of figurative use; all concordance lines are being used literally. Regarding the GENRE variable, the use of the [*es+schneeregnen*] construction is unsurprisingly and prodigiously informal. The vast majority of utterances (97%) belong to the genre type WEATHER TALK (cf. Table 70). There is only one concordance annotated as WEATHER REPORT, sourced from a weather forecast from a website dedicated to the weather in and around the city of *Düsseldorf*. Another two concordance lines belong to the LITERARY genre. Although WEATHER TALK is prevalent across constructions and precipitates, it appears the most striking here. To explicate, the other precipitates still encompass a good amount of formality within the genre type WEATHER TALK (in any of the constructions/precipitates). That means that sometimes it is not instantly apparent whether the concordance stems from an actual weather forecast or whether some "private" person has just been quoting the weather report on their personal blog or in a conversation on a message board.

Table 70. *Raw and relative requencies for the two variables TAM and GENRE for the* es*+PRED-Cx with the precipitate* Schneeregen *'sleet'*

Variable GENRE	Freq.	Relative frequency	Variable TAM	Freq.	Relative frequency
			PAST	31	**35%**
WEATHER TALK	85	**97%**	PRESENT	49	**56%**
WEATHER REPORT	1	**1%**	present_ FUTURE	4	**5%**
LITERARY	2	**2%**	FUTURE	2	**2%**
METEROLOGY	0	**0%**	IRREALIS	2	**2%**
other	0	**0%**	EXISTENTIAL/GENERIC	0	**0%**
Total	**88**	**100%**	**Total**	**88**	**100%**

The *es*+PRED-Cx for precipitate *Schneeregen* 'sleet' makes it clear, from the source websites URLs but also from the general tone of the utterances, that the weather talk is indeed mostly private and informal: a genuine proxy for informal spoken language. Correspondingly, the [*es+schneeregen*] corpus returns are highly imbued with stance, as we will see in subsection 3.3.3.4 on the MODIFIER variable.

3.3.3.2 The TAM (TENSE, ASPECT, MOOD) variable

Regarding the TAM variable, the majority of concordance lines is annotated as PRESENT (N=49; 56%), out of which a total of 42 furthermore express progressive aspect. Respecting quasi-auxiliaries, ten concordance lines focus on the incipient phase of the ME and contain variations

of *anfangen/beginnen zu* 'start/begin to'. Half of these occur in the progressive. The second most frequent tense is PAST (N=31; 35%). For PAST, as well, there are at least five concordance lines that describe the weather event as ongoing. Added up, more than 50% of the data mark progressive, which makes the [*es+schneeregnen*] construction highly immediate and puts the focus on the speaker's close experience of the ME. Eight concordance lines (9%) indicate future intent and there is not a single instance of EXISTENTIAL or GENERIC use.

Lastly, with respect to NEGATION, there are four concordance lines in total, which exhibit some sort of negation. Two concordance lines stand with *weder… noch* 'neither … nor', where it is neither *Schnee*/*Regen* nor *Schneeregen* that fell. One concordance says *mal nicht* 'for once not', where apparently there was no precipitation during the half-time of a soccer match, but it "sleeted" all through the rest of the game. And lastly, there is one concordance with *noch nicht* 'not yet', where sleet is supposedly expected.

3.3.3.3 The SPATIAL modifier variable

The frequencies of the SPATIAL modifier variable for the [*es+schneeregnen*] construction are given in Table 71 below. The majority of the data (N=51) does not contain any SPATIAL modification (ZERO). Where there is a SPATIAL modifier, it is mostly a LOCATION (40%) There are 25 concordance lines where a LOCATION given in the immediate context and another 12 where the location (L) is specified in the wider preceding or following context. Mostly the LOCATION takes the shape of deictic *hier* 'here', (N=11), another two are *bei uns* 'here' (lit. 'at ours']), or *draußen* 'outside' (N=9). These modifiers are the ones mostly found in the immediate context (or the KWIC). LOCATIONS in the left or right periphery more often take the shape of a proper name (e.g., 'I was born … in Berlin – it was a Wednesday and it sleeted') or a combination of adverbial plus proper place, e.g., *hier, ganz im Westen des Landes* 'here, in the very west of the country' or *vor der Sankt-Anna-Pfarrkirche im Lehel* 'in front of the St. Anna Church in Lehel'.

Table 71. *Raw and relative frequency of the* SPATIAL *modifier variable in the* es*+PRED-Cx and the preciptate* Schneeregen *'sleet'*

SPATIAL **modifier**	**Frequency**	**Relative frequency**
LOCATION	37	**42%**
GOAL	0	**0%**
SOURCE	2	**2%**
PATH	0	**0%**
ZERO	51	**58%**
Total	**90**	**100%**
COMBINATION L+S	1	1%

Lastly, there is one stand-alone SOURCE, as in (125), and one concordance where a LOCATION and a SOURCE are combined, as in (126).

(125) *Hier taut es inzwischen auch und* ***es schneeregnet von oben****.* (juwiswelt.blogspot.de 'juwi's-world')
'Meanwhile it is thawing here and **it is sleeting from above**.'

(126) ***Stadion*** *und das Warten* ***am Treffpunkt*** *sind grauenhaft: Mit den Füßen stecken alle in wässrigem Schneematsch,* ***von oben schneeregnet es*** *heftig.* (hinzundkunzt.de 'Hinz and Kunzt – a magazine for Hamburg, run in cooperation with the homeless')
'**The stadium** and waiting **at the meeting point** are terrible: Everybody's feet are stuck in the watery snowmud, **it is** severely **sleeting from above**.'

There are no goals and no paths, nor are there any other combinations of SPATIAL modifiers other than the one in (126) above.

3.3.3.4 The NON-SPATIAL modifier variable

The frequencies for the *es*+PRED-Cx and precipitate *Schneeregen* 'sleet', with respect to the presence of NON-SPATIAL modification, are given in Table 72. Since there are so few returns, the modifier types are not discussed individually in separate sub-sections (as done for previous precipitates and constructions), but, rather, are incorporated into the current section. The [*es*+*schneeregnet*] construction rarely appears without any modifiers (N=3, ZERO). There are 152 modifiers in total (144%), as in more than half of the data, at least two different types of modifier are combined, SPATIAL or NON-SPATIAL. Four different types of modifier are combined maximally, as in (128) below. The most frequent modifier is TEMPORAL and occurs in roughly two-thirds of the returns, followed by modifiers of STANCE, which are not quite as frequent but are also the second most frequent stand-alone modifier (N=13). With respect to their semantic categories, TEMPORAL modifiers contain mostly "deictics", i.e., adverbials, such as *heute* 'today',

jetzt 'now', *mittlerweile* 'meanwhile', but also nominals, such as *im Moment* 'at the moment', which also has the highest incidence where it is the only modifier (N=14). Modifiers of TIME also express "frequency" (*täglich* 'daily'), "duration" (*seit Stunden* 'for hours'), or containan "exact date", like *Dienstag*, *19. März 2013* 'Tuesday, March 19,

Table 72. *Raw and relative frequencies for the NON-SPATIAL modifier variable within the* es*+PRED-Cx and the precipitate* Schneeregen *'sleet'*

NON-SPATIAL **modifier**	**Frequency**	**Relative frequency** / 88
TEMPORAL	55	**62.5%**
QUANTITY	8	**9%**
QUALITY	11	**12.5%**
STANCE	53	**60%**
Total	**127**	**144%**
ZERO	3	3%
COMBINATIONS	49	56%
SPATIAL modifier	25	28%

Regarding the semantics of STANCE, speakers mostly express a negative attitude towards the ME by using words which exhibit negative semantic prosody or are more explicit unfavourable expressions of stance. Accordingly, there are simple declarations, like *leider* 'unfortunately', but also the more emphatic *Furchtbar!* 'awful!' or *Mistwetter!* 'crap weather', both of which accompanied by an exclamation mark. These are also stand-alone modifiers. STANCE, however, can also take the shape of a litany of adversities, topped off with a spell of bad weather, sleet as in example (127). There is only one concordance where the event of *Schneeregen* is perceived as neutral or positive, which is the description of the opening night of a Christmas market where sleet falling is deemed to be better than no snow at all. Otherwise, sleet is only perceived to be positive when it is absent. Example (128) encompasses another expression of STANCE (*nicht so schön* 'not so nice') and is one of two concordance lines where a maximum of four modifiers are combined: LOCATION (*hier* 'here')+TIME (*jetzt noch* 'now still')+QUALITY (*zum Teil* / partly)+STANCE (*nicht so schön* 'not so nice'). There is no concordance that combines all five modifiers.

(127) *Heute war mal einer der schlechten Tage. Alles ging schief, alle haben genervt, meinem Freund ging es nicht so gut, alles kam auf einmal, echt ein scheiss Tag. ich bin heulend von Köln zurück nach Siegen gefahren,* ***dann hat es auch noch angefangen zu schneeregnen***. (*krebs-kompass.de* 'cancer-compass')
'Today was one of the bad days. Everything went wrong, everyone was annoying, my boyfriend wasn't doing so well, everything came at once, really a shit day. I drove back from Cologne to Siegen crying, **and then on top of that it started sleeting**.'

(128) *Hier hat es die ganze letzte nacht geschneit und zum teil* ***schneeregnet es*** *jetzt noch…****nicht so schön***. (amalielovesdenmark.blogspot.ch)
'Here, it snowed all of last night and partly **it is** still **sleeting**…**not so nice**.'

The least frequent modifiers are QUALITY (N=11) and QUANTITY (N=8). The QUANTITY modifier mainly appears in combination with the TEMPORAL modifier and four times also with STANCE. Five modifiers refer to small quantities like *ein bisschen* 'a little/a bit', *ein wenig* 'a little/some' and ein paar Flöckchen 'a few flakes$_{DIM}$'. One token refers to the MANNER or a large QUANTITY (colloq. *was nur ging* 'what was possible'). The QUALITY modifiers are a bit more varied referring mostly to the "intensity" and "manner" (*in Strömen* 'pouring' lit. 'in streams'; *wie aus Kübeln* 'buckets', *heftig* 'heavily/severely', *fröhlich vor sich hin* 'merrily away'. Again, the wider preceding or following context was included, so that a QUALITY modifier can specify the ME in a preceding or subsequent sentence, as in (129), where the QUALITY is given in the second sentence in the form of a relative clause, which strictly speaking modifies the noun *Niederschlag* 'precipitation', which, of course, refers back, to the sleet ME of the first sentence.

(129) *Zum Ende des Umzugs* ***fing es an zu schneeregnen****.* ***So ein fieser Niederschlag*** *aus dicken nassen Flocken,* ***der einen in Sekundenschnelle durchweicht***. (kggp.de *Keine Gewalt gegen Polizei* 'no-violence-against-police')
'At the end of the parade **it started to sleet**. **Such a nasty precipitation** of thick wet flakes **that soaks you in seconds**.'

Example (129) is furthermore one of ten that pinpoint the incipient phase of the ME by adding the quasi-auxiliary *anfangen* 'start' (N=8) or *beginnen* 'begin' (N=2). There is not a single concordance line where the terminative phase of the event is in focus. The next subsection compares the results of the three constructions and their interaction with the various variables discussed up to this point.

*3.3.4 Distribution and frequencies of variables across the three constructions (ARG+*fallen*, ARG+*geben*, es+PRED) with the precipitate* Schneeregen *'sleet'*

This section surveys the aggregate frequencies of the three constructions in combination with the precipitate *Schneeregen* 'sleet' and their interactions with the different variables as discussed in the sections above and compares their frequencies and usage against one another. The order given follows earlier presentation orders starting with the GENRE variable followed by the TAM, SPATIAL, and NON-SPATIAL modifier variables. The aggregate frequencies include the

returns for the [*es+schneeregnen*] construction, keeping in mind that the total of the *es*+PRED-CX data comes to only 88 concordance lines as opposed to 200 concordance lines for each of the two ARG-constructions. Comparatively speaking, the novel verb form, *schneeregnen*, is dispreferred.

3.3.4.1 USAGE (LITERAL VS FIGURATIVE) AND GENRE VARIABLES ACROSS THE THREE CONSTRUCTIONS

All three constructions based on the precipitate *Schneeregen* 'sleet' occur exclusively with literal meaning. There is not a single instance of figurative usage in all of the returned concordance lines. The distribution and frequencies of the GENRE variable by construction are given in Table 73 and are visualized in Figure 22. Despite the overall preponderance of genre WEATHER TALK, with two-thirds of the corpus returns for precipitate *Schneeregen* 'sleet', there is a clear bias of WEATHER TALK attracting the [*Schneeregen+geben*] construction over either the [*Schneeregen+fallen*] construction and the [*es+schneeregnen*] construction. In turn, WEATHER REPORTS chiefly attract the *Schneeregen+fallen*-Cx for WEATHER REPORTS overall (70%) but also within the [*Schneeregen+fallen*] construction itself, where more than half of the concordance lines are coded as WEATHER REPORTS. The less frequent [*es+schneeregnen*] construction only contains a single utterance coded as WEATHER REPORT. Furthermore, it comprises roughly as many concordance lines coded as WEATHER TALK as the [*Schneeregen+fallen*] construction (N=85 vs. N=83; each 26%). Even so, the [*es+schneeregnen*] construction only contains two more concordance lines in total, both of which are from LITERARY sources, but not a single utterance from the METEOROLOGICAL genre. The latter two genre types are few and far between, particulary METEOROLOGY, which only accounts for three tokens in total (0.06%). The LITERARY genre, at least, amounts to 14 tokens (2.9%), which mainly belong to the [*Schneeregen+fallen*] construction.

Table 73. *Raw and relative frequencies of the GENRE variable across the three constructions (ARG+*fallen*, ARG-*geben*, es+PRED) and precipitate* Schneeregen *'sleet', (abbreviated as SR)*

Construction / GENRE	*SR+fallen*	*SR+geben*	*es+schneeregnen*	Total	Relative Frequency
WEATHER TALK	83	154	85	**322**	**66%**
WEATHER REPORT	105	43	1	**149**	**30.5%**
METEOROLOGY	2	1	0	**3**	**0.5%**
LITERARY	10	2	2	**14**	**3%**
Total	**200**	**200**	**88**	**488**	**100%**

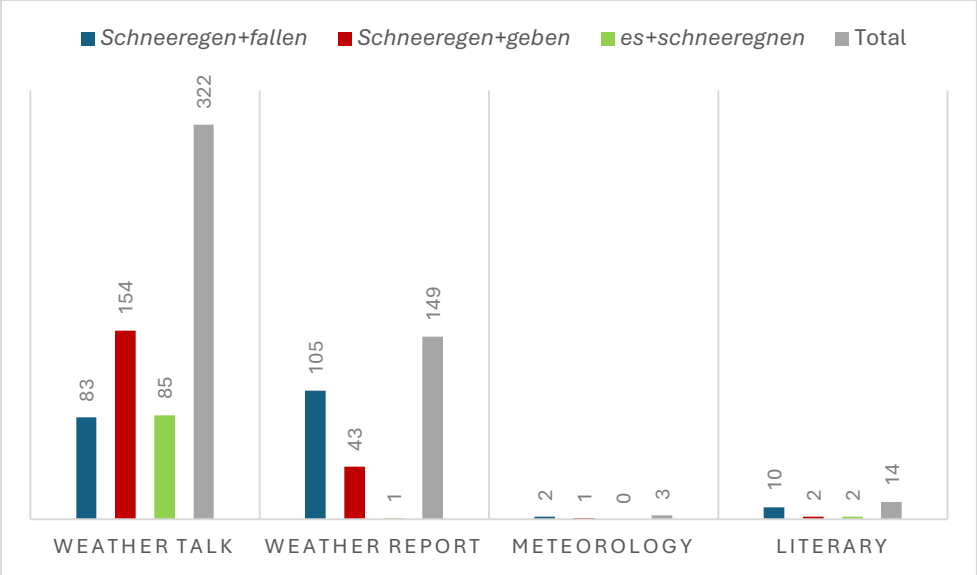

Figure 22. *Distribution and absolute frequencies of the* GENRE *variable across the three constructions (ARG-*fallen*, ARG-*geben*,* es*+PRED) and the precipitate* Schneeregen *'sleet'*

In sum, the genre type WEATHER TALK is clearly the most frequent genre overall and attracts the [*Schneeregen+geben*] construction and the [*es+schneeregnen*] construction in particular. By contrast, genre type WEATHER REPORT predominantly attracts the [*Schneeregen+fallen*] construction. The genre types LITERARY and METEOROLOGY are infrequent. Nevertheless, the [*Schneeregen+fallen*] construction, moderately, is attracted by the LITERARY genre to some extent.

3.3.4.2 The TAM (TENSE, ASPECT, MOOD) variable across the three constructions

This section surveys the distribution and frequencies of the TAM variable across the three constructions (ARG+*fallen* ARG+*geben*, and the *es*+PRED-Cx) in combination with the precipitate *Schneeregen* 'sleet'. The absolute frequencies of the TAM variable for each construction are given in Table 74 and visualized in Figure 23. The percentages of how much each construction contributes to a particular level of tense are shown in Figure 24. Since the TAM levels IRREALIS and FUTURE have very low incidences, they were collapsed into one category "future intent" (FUTURE_intent) together with tokens coded as present_FUTURE. These three TAM levels accumulated amount to the same frequency as PAST, i.e., 180 tokens (or 37% each). This makes PAST tense and FUTURE_intent also the highest contributing levels of tense overall. In stark contrast to these frequencies are the EXISTENTIAL_GENERIC tokens, which barely amount to 2% of

the data overall. Covering the middle ground are concordance lines in the PRESENT (and present progressive), roughly responsible for one-fourth of the returns. Despite the *es+schneeregnen*-Cx's overall lower incidence, it still accounts for 41% of the PRESENT tense data. This is not quite as much as [*Schneeregen+fallen*] (45%) but much more than [*Schneeregen+geben*] (14%). Moreover, 35% of the [*es+ schneeregnen*] constructions exhibit progressive aspect.

Table 74. *Raw and relative frequencies of the TAM variable across the three constructions (ARG+*fallen*, ARG+*geben*, es+PRED) and the precipitate* Schneeregen *'sleet' (abbreviated as SR)*

Construction/ TAM	*SR+fallen*	*SR+geben*	*es+schneeregnen*	Total	Relative Frequency /488
PAST	44	105	31	**180**	**37%**
PRESENT	53	17	49	**119**	**24%**
present_FUTUR	90	65	4	**159**	**33%**
FUTURE	7	10	2	**19**	**4%**
IRREALIS	0	0	2	**2**	**0%**
EXIS_GEN	6	3	0	**9**	**2%**
Total	**200**	**200**	**88**	**488**	**100%**

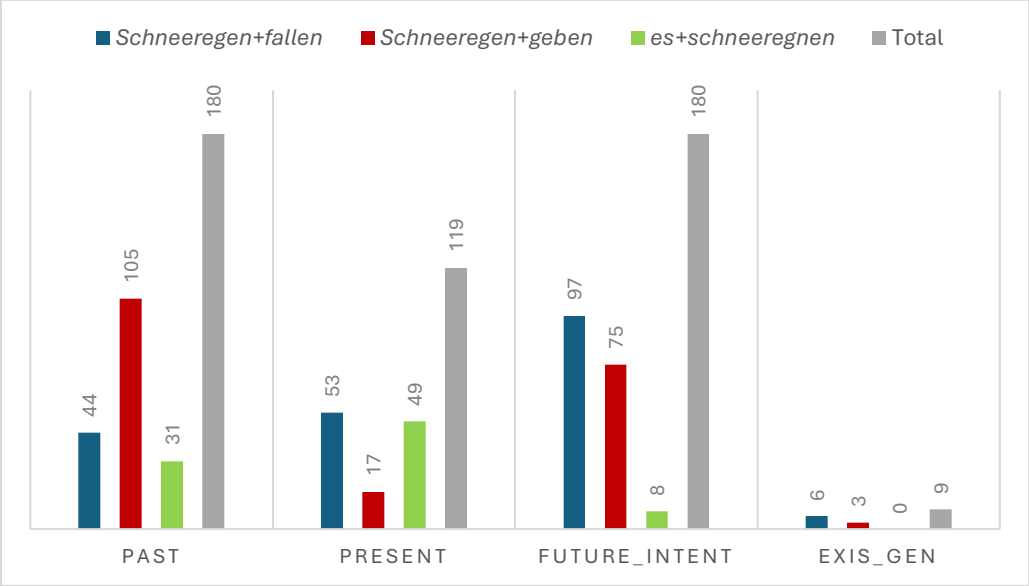

Figure 23. *Distribution of absolute frequencies of the TAM variable across the three constructions (ARG-*fallen*, ARG-*geben*, es+PRED) and the precipitate* Schneeregen *'sleet'*

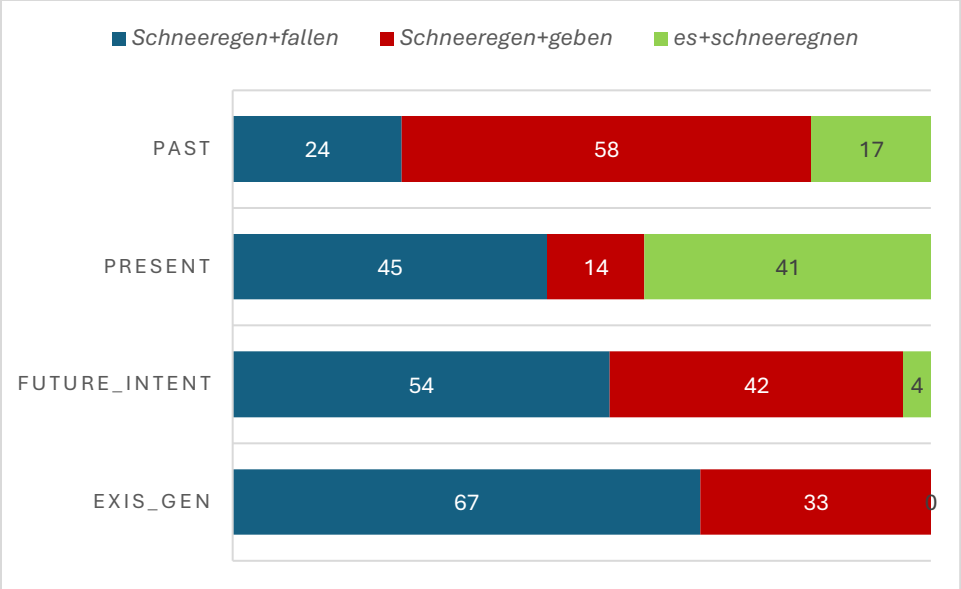

Figure 24. *Relative frequencies of the* TAM *variable for the precipitate* Schneeregen *'sleet' by construction (ARG+*fallen*, ARG+*geben*, es+PRED) with levels present_* FUTURE*, FUTURE, and* IRREALIS *collapsed under* FUTURE*_intent*

Within the PAST tense data sets, the [*Schneeregen+geben*] construction is the main contributor and within the FUTURE_intent data sets, the [*Schneeregen+fallen*] construction is responsible for the bulk of returns, which neatly goes together with the preponderance of WEATHER REPORTS (or, rather, future weather predictions). In comparison, the [*Schneeregen+geben*] construction contributes a lower, but still fairly high incidence of FUTURE_intent tokens, but a much lower incidence of PRESENT tense tokens, which are mostly provided by the [*Schneeregen+geben*] construction and the [*es+schneeregnen*] construction, despite the overall low incidence of the latter. The [*es+schneeregnen*] construction, furthermore, contributes 17% of the PAST tense tokens; this is 35% of PAST tense within its own data set (cf. Table 70). As we will see in Section 3.5, the tendency to speak about MEs in the past tense is fairly evenly distributed between the precipitate types, but it is not as pronounced within the precipitate set at hand: the *Schneeregen* 'sleet' data sets contribute the lowest percentage of PAST tense instances overall. Once the levels present_FUTURE, FUTURE, and IRREALIS are collapsed into one category (FUTURE_intent), PAST and FUTURE events are referenced the exact same amount frequencywise.

3.3.4.3 The SPATIAL *modifier variable across the three constructions*

This section examines the distribution and frequencies of the SPATIAL modifier variable across the three constructions (ARG+*fallen* ARG+*geben*, es+PRED) in combination with precipitate

Schneeregen 'sleet'. The relative and absolute frequencies of the SPATIAL modification for each construction are given in Table 75 and the absolute frequencies are visualized in Figure 25. There is no figure visualizing the relative frequencies of each SPATIAL modifier, since the distribution of all three constructions is chiefly happening between LOCATIONS and ZERO modification and the more peripheral modifiers occur predominantly within the [*Schneeregen+fallen*] data sets. Overall, static LOCATIONS are the most frequently occurring SPATIAL modifiers with more than half of the corpus returns for the *Schneeregen* 'sleet' precipitate. The second highest incidence is that of ZERO SPATIAL mentions in more than one-third of the data. Regarding the less frequent SPATIAL modifiers, GOALS, unsurprisingly, have the highest incidence, albeit only by a small margin and only two tokens more than SOURCES. PATHS are uncommon overall.

Table 75. *Raw and relative frequencies of the SPATIAL modifier variable across the three constructions (ARG+*fallen*, ARG+*geben*, es+PRED) and the precipitate* Schneeregen *'sleet' (abbreviated as SR)*

Construction / SPATIAL Modifier	***SR+fallen***	***SR+geben***	***es+schneeregnen***	**Total**	**Relative frequency /488**
LOCATION	110	109	37	**256**	**52%**
GOAL	25	3	0	**28**	**6%**
SOURCE	20	4	2	**26**	**5%**
PATH	3	1	0	**4**	**1%**
ZERO	50	86	51	**187**	**38%**
Total	**208**	**203**	**90**	**501**	**102%**
COMBINATIONS	8	3	1	12	2.5%

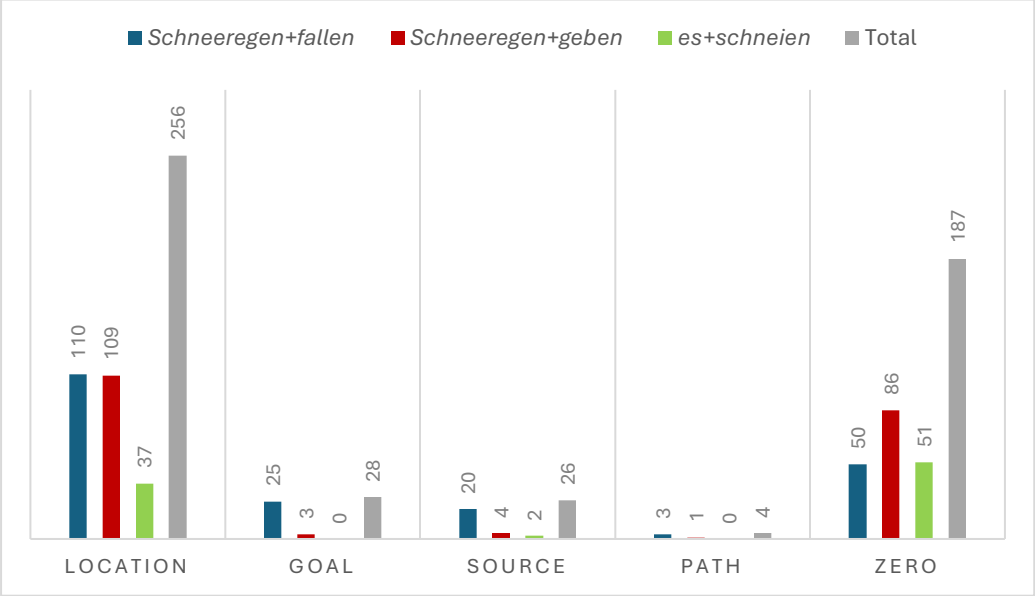

Figure 25*. Distribution and absolute frequencies of the SPATIAL modification variable across the three constructions (ARG-*fallen*, ARG-*geben*, es+PRED) and precipitate* Schneeregen *'sleet'*

Semantically and syntactically, the SPATIAL modifiers differ in that both ARG-Cxx more closely resemble Liégeois's (2021) findings for weather reports, i.e., they are predominantly nominal. Additionally, the [*Schneeregen+fallen*] construction contains mainly references to altitude, either by using numerals (*8000m*) or lexical items (*im Bergland* 'in the mountainous regions') where especially the former often come in the shape of a GOAL. As we have already seen with the [ARG+*fallen*] constructions with precipitates *Regen* 'rain' and *Schnee* 'snow', the same construction with precipitate *Schneeregen* 'sleet' contains the highest incidence of GOALS, SOURCES, and PATHS, which often occur in some combination. The frequent mention of these peripheral modifiers makes the [*Schneeregen+fallen*] construction more physical and descriptive, while the [*Schneeregen+geben*] construction appears more circumstantial and abstract. In line with the high incidence of GOALS, SOURCES, and PATHS , the [*Schneeregen+fallen*] construction encompasses twice as many combinations as the other two constructions combined. Syntactically, apart from the [*Schneeregen+geben*] construction containing many nominal SPATIAL modifiers, it also contains many adverbial deictics, the most frequent types being *hier*/*bei uns* 'here' (N=14).

3.3.4.4 The NON-SPATIAL modifier variable across the three constructions

This section discusses the distribution and frequencies of the NON-SPATIAL modifier variable across the three constructions (ARG+*fallen,* ARG+*geben*, and the es+PRED-Cx) in combination with the precipitate *Schneeregen* 'sleet'. The absolute frequencies of the NON-SPATIAL modifier variable for each construction are given in Table 76 and visualized in Figure 26. The relative frequencies of variable MODIFIER within the *Schneeregen* ('sleet') data are shown in Figure 27. As we have seen with other constructions and precipitates, the three constructions and precipitate *Schneeregen* 'sleet' contain a similar overall number of NON-SPATIAL modifiers. That means, the ARG-Cxx encompass 221 and 278 NON-SPATIAL modifiers within 400 lines of data, while the *es*+PRED-Cx encompasses 127 NON-SPATIAL modifiers within 88 lines of data. Adjusted to 200 concordance lines, the NON-SPATIAL modifier total would arrive at a slightly lower frequency than the ARG-Cxx. TEMPORAL modifiers occur in more than two-thirds of the returns and have the highest incidence within each construction type. The second most frequent NON-SPATIAL modifier, STANCE, follows by a wide margin with less than half as many instances (N=153; 31%). Modifiers of QUALITY are the third most frequent, again with roughly half the amount of STANCE

tokens, and the lowest incidence is provided by QUANTITY modifiers, which only add up to 10% of the data (N=50).

Table 76. *Distribution and absolute frequencies of variable MODIFIER across the three constructions (ARG+*fallen*, ARG+*geben*, es+PRED) and precipitate* Schneeregen *'sleet' (abbreviated as SR)*

Construction / NON-SPATIAL modifier	***SR+fallen***	***SR+geben***	***es+schneeregnen***	**Total**	**Relative frequency / 488**
TEMPORAL	124	158	55	**337**	**69%**
QUANTITY	26	16	8	**50**	**10%**
QUALITY	41	34	11	**86**	**17%**
STANCE	30	70	53	**153**	**31%**
Total	**221**	**278**	**127**	**626**	**127%**
ZERO	5	4	3	12	2%
COMBINATIONS	137	139	49	325	66.5%
SPATIAL modifier	151	96	25	272	56%

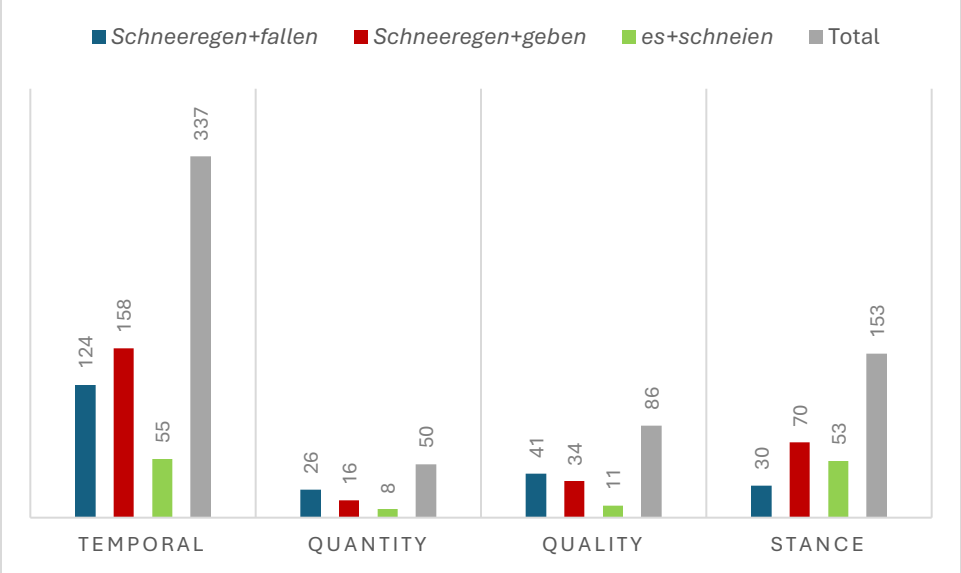

Figure 26. *Distribution and absolute frequencies of the NON-SPATIAL modifier variable across the three constructions (ARG-*fallen*, ARG-*geben*, es+PRED) with the precipitate* Schneeregen *'sleet'*

The relative frequencies that each construction contributes to the individual levels of NON-SPATIAL modifier in combination with precipitate *Schneeregen* 'sleet' are given in Figure 27. By construction, [*Schneeregen+geben*] accounts for almost half of the TEMPORAL modifiers, followed by [*Schneeregen+fallen*] with 10 percentage points less, and the [*es+schneeregnen*] construction makes up the remaining 16%. In proportion to each construction's number of concordance lines, the three constructions contain similar incidences of TEMPORAL modifiers; in fact, the

[*Schneeregen+fallen*] corpus returns contain 62% of TEMPORAL modifiers and the [*es+schneeregnen*] corpus returns contain 62.5% of TEMPORAL modifiers.[91]

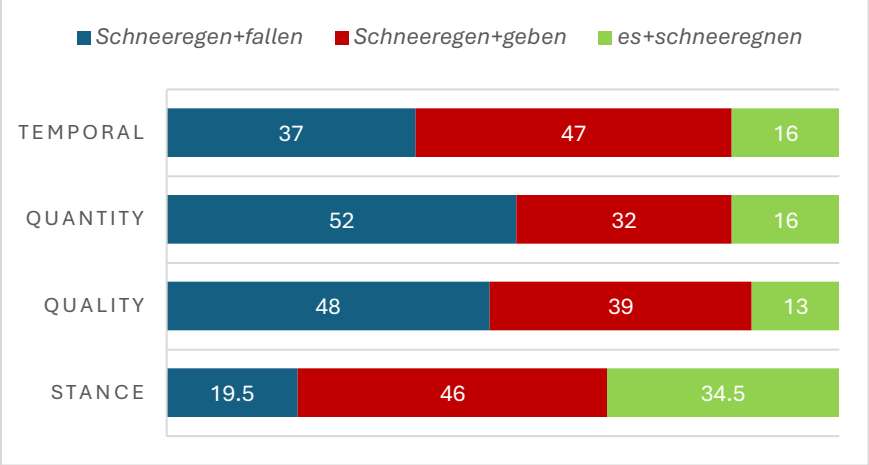

Figure 27. *Percentual proportion each construction (ARG-*fallen*, ARG-*geben*, es+PRED-Cx) contributes to the total of each level of the NON-SPATIAL modifier variable with the precipitate* Schneeregen *'sleet'*

The NON-SPATIAL modifiers with lower incidences, i.e., QUANTITY and QUALITY, are predominantly present in the [*Schneeregen+fallen*] construction, which accounts for more than half of the QUANTITY modifiers and almost half of the QUALITY modifiers respectively. [*Schneeregen+geben*] still contains a comparatively high percentage of QUALITY and less of the QUANTITY modifiers. The [*es+schneien*] construction contributes the same proportion of QUANTITY tokens as TEMPORAL tokens (each 16%) but a lower proportion of QUALITY tokens (12%) compared to 5% of QUANTITY tokens and 7% QUALITY of tokens within its own data set (cf. Table 72). Finally, the *Schneeregen+geben*-Cx contributes almost half of the STANCE tokens, followed by the [*es+schneien*] construction with 35%. Both constructions also contain high amount of stancetaking within their own data sets. More than one-third of the [*Schnee+geben*] returns contain expressions of STANCE, but even more strikingly, more than half of the [*es+schneien*] returns do so. In contrast, only 15% of the *Schneeregen+fallen* returns contain expressions of STANCE.

[91] Number of TEMPORAL modifiers divided by the total of concordance lines per construction, i.e., 55/88=62.5 (*es+schneeregnen*) and 124/200=62 (*Schneeregen+fallen*).

3.3.5 Summary of Schneeregen *'sleet' constructions*

The preceding quantitative and qualitative analysis of MEs employing the three constructions ARG+*fallen*, ARG+*geben*, and *es*+PRED in combination with precipitate *Schneeregen* 'sleet' has provided a comprehensive account of the constructional preferences of each of the three constructions and their interaction with a number of variables. While there is overlap between the three constructions, each construction was also found to attract different variables to varying degrees. The results also show a departure from the two more "canonical" precipitates *Regen* 'rain' and *Schnee* 'snow'. First of all, this finding is most evident in the very informal and tentative usage of the *es*+PRED-Cx, which is has not proven itself frequent enough in language use to receive an entry in the *Duden* and does not seem to be equally acceptable by German speakers (cf. Heinrich 2017). Secondly, the ARG-Cxx's propensity to occur with a high incidence of genre type WEATHER REPORTS and the corresponding tendency to inflect a large number of utterances for "future intent" with respect to the TAM variable is the most pronounced within the *Schneeregen* 'sleet' precipitate data sets.

A Pearson's chi-square test was carried out for Table 73 (GENRE variable), Table 74 (TAM variable), Table 75 (SPATIAL) and Table 76 (NON-SPATIAL).

For the first three mentioned tables, the frequencies are too low for a reliable application of the chi-squared test for a significant distribution. Nevertheless, Table 76 yielded a highly significant result (Table 76: X-squared = 42.272, df = 6, p-value = 1.625e-07). For Table 75, where the peripheral spatial modifiers (goal, source, and path) are combined, the distribution is highly significant (X-squared = 63.183, df = 4, p-value = 6.211e-13). As with the *Regen* 'rain' and *Schnee* 'snow' Cxx, the peripheral paths are much more frequent than expected where the ARG+*fallen* Cx is concerned and much less frequent where the ARG-*geben* Cx is concerned. The more general observations are summarized in the following:

The [ARG+*fallen*] construction in combination with precipitate *Schneeregen* 'sleet' is used exclusively in the literal sense. Regarding the GENRE variable, it is predominantly attracted by WEATHER REPORTS (with more than half of the data), followed by WEATHER TALK, but also contains some utterances from LITERARY sources. Tensewise, this construction is firmly rooted in the realm of "future intent" with approximately half of the corpus returns projecting imminent MEs, which ties in neatly with the preponderance of WEATHER REPORTS. The remainder of the data can roughly be divided into PRESENT and PAST tense events, with the former having the edge over

the latter, which is, furthermore, a "monolith" of preterite tense. Regarding variable SPATIAL modification, half of the returns contain a LOCATION, which immediately modifies the ME, one-fourth of the data does not contain any SPATIAL mentions whatsoever, and the remainder can roughly be split into GOALS, SOURCES, and PATHS, which is a remarkably high number for SPATIAL mentions that do not denote LOCATIONS. Regarding SPATIAL and NON-SPATIAL modifiers, LOCATIVES are even more frequent than modifiers of TIME and are present in three-fourths of the returns. Yet, at least one TEMPORAL modifier still is present in around two-thirds of the data. The other modifiers trail behind; having said that, QUALITY modifiers are moderately frequent and show up in one-fourth of the concordance lines and mostly denote "lightness". Modifications of QUANTITY and STANCE, on the other hand, are fairly infrequent. QUANTITY mentions tend to point to small quantities; the most frequent type is *etwas* 'a little'.

The [ARG+*geben*] construction in combination with precipitate *Schneeregen* 'sleet', too, is a literal construction with no room for figurative use. Regarding the GENRE variable, the [*Schneeregen*+*geben*] construction can be roughly divided into three-fourths WEATHER TALK, which may indicate a higher informality of speech, and one-fourth WEATHER REPORTS with almost no effect of genre levels LITERARY and METEOROLOGICAL. More than half of the data is in PAST tense with some variation of *Perfekt* and past perfect within the PREC1 order but simple past only in the PREC2 order. The second largest level of the TAM variable, with more than one-third of the data, is "future intent" where almost all of the PREC1 order contains a modal verb. The SPATIAL modifier variable is mostly distributed between ZERO path and LOCATIONS, with very few LOCATIONS not being in the immediate context of the ME. Semantically, the SPATIAL modifiers are chiefly "proper locations" followed by "deictics". Regarding the NON-SPATIAL modifier variable, about four-fifths of the data contain at least one modifier of TIME. Moreover, TEMPORAL modifiers occur with many types and tokens; the most frequent type is adverbial deictic *heute* 'today'. STANCE is frequently expressed within the [*Schneeregen*+*geben*] construction, i.e., in more than one-third of the returns and is almost invariably negative. Modifiers of QUALITY have little bearing, however, they tend to focus on the severity or intensity of the ME. Modifiers of QUANTITY are very infrequent, but where they do occur, they indicate small quantities over large quantities.

The [*es*+PRED] construction, too, is used exclusively in a literal way. As pointed out earlier, the use of the construction in combination with the precipitate *Schneeregen* 'sleet' is marked in that it has not yet entered the standard language. This observation appears to correspond to its interaction with a particular set of extra-propositional variables. As opposed to

the ARG-Cxx, it almost exclusively attracts the highly informal genre type WEATHER TALK as opposed to WEATHER REPORTS. The informality is mirrored in its immediacy and use in chiefly PRESENT tense, which is, furthermore, commonly marked for progressive aspect. PAST tense use is secondary, but contains some variation with *Perfekt* and past perfect use, and also progressives. Moreover, there are ten quasi-auxiliaries, which exclusively focus on the incipient phase of the ME. Regarding NON-SPATIAL modifiers, more than half of the data contain expressions of STANCE, which, in turn, are almost entirely negative. The high incidence of STANCE expressions additionally supports the claim of informality. Only modifiers of TIME are slightly more frequent than modifiers of STANCE. SPATIAL modifiers are less frequent than modifiers of TIME and STANCE and mostly contain adverbial deictics. However, it is more likely for an utterance to contain ZERO spatial modification. Modifiers of QUALITY and especially QUANTITY are relatively rare. In accordance with this finding, paths do not occur in more than half of the concordance lines. The following section constitutes the final univariate survey of the final precipitate *Hagel* 'hail'.

3.4 Constructions with the precipitate *Hagel* 'hail'

This section examines the *es*+PRED-Cx as well as the two ARG-Cxx ([ARG+*fallen*] and [ARG+*geben*]) with the *Hagel* 'hail precipitate. Hail is a generally a less frequently occurring ME, which needs specific conditions to form and is also more likely to happen in certain regions (the German south more than the north) at certain times of the year (in Germany particularly the month of August). Meteorologically, hail forms in thunderstorm updrafts, which means it is a byproduct or co-ME of a different, non-precipitate ME altogether.[92] What the precipitate lacks in naturally occurring frequency, it makes up for in salience. Since hail can cause serious damage and even be lethal, it often associates with negative semantic prosody and occurs in discourse that does not specifically focus on the ME itself, but its effects, e.g., insurance texts that deal with hail damage. Moreover, its salience and destructive nature make hail a prime source for figures of speech; accordingly, the *es*+PRED-Cx returned an overwhelming amount of

[92] Raindrops are carried upward by thunderstorm updrafts above freezing level, which turns them into hailstones. These hailstones can grow even bigger when they collide with other water drops until they finally become too heavy for the updrafts and fall to the ground. (Source: "Hail Basics – Severe Weather 101" nssl.noaa.gov, last accessed December 31, 2023).

metaphorical utterances, which lead to a much lower incidence of analyzable concordance lines than anticipated (N=405 instead of the expected approximately N=600). The same is not true for the two ARG-Cxx, however, which indicates that the precipitates do not equally associate with the [*es*+PRED] construction. The distributions and frequencies for the *Hagel* 'hail' precipitate in general turn out to deviate in some regard from the other precipitates. Having said that, we need to keep in mind that the frequencies of the *Hagel* 'hail' precipitate data sets are much lower than those for the other precipitates; this is especially pertinent to the aggregate frequencies and comparison of the three constructions. As a consequence, the following subsections are predominantly indicative of the two ARG-Cxx.

3.4.1 The argument construction [Hagel+fallen] *'hail +fall'*

The queries in (xxv) and (xxvi) were carried out to obtain examples containing the [*Hagel*+ *fallen*] construction. The first query returned 330 concordance lines (22 concordance lines had to be removed as false returns or duplicates). The second query returned 266 results (67 concordance lines had to be removed as false returns).

(xxv) [word="Hagel"] [ ] {0,8} [lemma="fallen"] within <s/> (N=330)
(xxvi) [lemma="fallen"] [ ] {0,8} [word="Hagel"] within <s/> (N=266)

From the set of 308 and 199, each one-hundred concordance lines were selected randomly to obtain 200 utterances, which were coded for the variables surveyed below.

3.4.1.1 USAGE (LITERAL VS FIGURATIVE) and GENRE variables

The [*Hagel*+*fallen*] construction overwhelmingly gets used literally (190 out of 200 concordance lines) as shown in Table 77.

Table 77. *Frequencies of literal and figurative USAGE of the ARG+*fallen *construction with the precipitate* Hagel *'hail'*

Word order / usage	**PREC1** *Hagel+fallen*	**PREC2** *fallen+Hagel*	**Total**
Literal	95	95	**190**
Figurative	5	5	**10**
Total	**100**	**100**	**200**

There are ten instances of figurative use, often metaphorical, as in (202). The above queries returned a total of 566 concordance lines from the *deTenTen13*, which is a much lower overall return than for the [*es+hageln*] constructions, which returned roughly 38 times as many concordance lines (N=21,435). Nevertheless, the above queries, especially (xxv), returned a high amount of actual, literal weather events, while the [*es+hageln*] constructions are used predominantly in a figurative manner. The false returns from queries (xxv) and (xxvi), which were removed, were, however, not metaphorical uses, but ones where the predicate *fallen* 'fall' relates to a different argument.

(202) *(Eltern) (…)* ***lassen*** *gerne mal* ***ein*** [sic] ***Hagel von Ermahnungen*** *auf den Tisch und die Kinderköpfe* ***fallen****.* (difoer.de)
'(Parents) like to **let a hail of admonitions fall** onto the table and their children's heads.'

The distribution and frequencies of the GENRE variable within the [*Hagel+fallen*] construction are given in Table 78. WEATHER TALK is the most frequent type of genre with about one-third of the returns (N=65, 34%), and as exemplified in (203).

(203) *Aber leider war es mir nicht vergönnt, dieses Jahr auch nur überhaupt einmal* ***Hagel fallen*** *zu sehen.* (scvswitzerland.ch)
'But, unfortunately, I wasn't lucky enough to see any **hail fall** at all this year.'

The second most frequent, with about two-thirds as many tokens as WEATHER TALK (N=46; 24%), are WEATHER REPORTS, followed closely by utterances from LITERARY sources (N=42; 22%). METEOROLOGY is the least frequent genre type with roughly one-fifth of the returns (N=37; 19%). The ten concordance lines that make up the figurative category mostly belong to the LITERARY genre. Nine are either from religious texts or fiction, and one is taken from a political excerpt regarding animal rights.

Table 78. *Raw and relative frequencies of the GENRE variable for the ARG+fallen-Cx and the precipitate* Hagel *'hail'*

Word order/	**PREC1**	**PREC2**	**Total**	**Relative frequency**
GENRE	***Hagel+fallen***	***fallen+Hagel***		**/190***
WEATHER TALK	26	39	**65**	**34%**
WEATHER REPORT	19	27	**46**	**24%**
METEOROLOGY	22	15	**37**	**19%**
LITERARY	28	14	**42**	**22%**
Total	**95**	**95**	**190**	**100%**
FIGURATIVE	5	5	10	**5%

***only literal instances, ** divided by 200 (figurative instances included)

In brief, while WEATHER TALK is the most frequent type of genre by a modest margin, genre types WEATHER REPORTS, METEOROLOGY, and LITERARY are close to each other with nine tokens difference at the maximum. Across the board (meaning across all constructions and precipitates examined in this dissertation), the [*Hagel+fallen*] construction returns the highest incidence of LITERARY genre type utterances. Including the non-literal utterances, more than one in four concordance lines can be categorized as LITERARY. The sources of these concordance lines are fairly evenly distributed between prose fiction as in (204) and religious texts as in (205) below; additionally, one concordance refers to mythology and two are from poetry.

(204) ***Der <u>Hagel fiel</u>*** *so dicht, daß sich die Dämmerung von dem vielen Weiß, das den Grund bedeckte, zu erhellen begann.* (gutenberg.de)[93]
'**The <u>hail fell</u>** that densely that the dusk/dawn began to lighten up from all the white that covered the ground.'

(205) *Nun sagte der HERR zu Mose: "Streck deine Hand zum Himmel aus, damit in ganz Ägypten **<u>Hagel fällt</u>**, auf die Menschen, die Tiere und alle Pflanzen auf dem Feld."*
'Then the LORD said to Moses, "Stretch out your hand toward the sky so that **<u>hail falls</u>**
all over Egypt - on men and animals and all the plants in the fields."' (erneuertdurchseinwort.de – 'regeneratedbyhisword')

Overall, the differences between types of genre are not as stark as we have seen for other precipitates.

3.4.1.2 The TAM (TENSE, ASPECT, MOOD) variable

The frequencies of the [*Hagel+fallen*] construction and their distribution by TAM level are given in Table 79. The highest frequency, amounting to more than half of the returns, is covered by PAST tense events (N=105; 55%). Second most frequent is the EXISTENTIAL/GENERIC use with roughly half as many concordance lines (N=51; 28%). The remainder are made up of concordance lines in PRESENT tense (N=20; 10%), and instance of "future intent" (N=12, or 6%), i.e., tokens of present_FUTURE, IRREALIS, and FUTURE. The periphrastic future tense contributes more than half of these instances. Out of the non-literal usages (mostly metaphors), five are coded as PAST, three as EXISTENTIAL/GENERIC and two as PRESENT tense.

[93] The Gutenberg document is a historical novel called *Der Mann im Salz: Roman aus dem Anfang des 17. Jahrhunderts* ('Sealed in Salt: A Novel from the Beginning of the 17th Century') by German author Ludwig Ganghofer and was originally published in 1906.

Table 79. *Raw and relative frequencies of the* TAM *variable for the ARG+*fallen*-Cx and the precipitate* Hagel *'hail'*

Word order/ TENSE	PREC1 *Hagel+fallen*	PREC2 *fallen+Hagel*	Total	Relative frequency /190*
PAST	41	64	**105**	**55%**
PRESENT	12	8	**20**	**10%**
present_ FUTURE	3	1	**4**	**2%**
FUTURE	7	0	**7**	**4%**
IRREALIS	1	0	**1**	**5%**
EXIS_GEN	31	22	**53**	**28%**
Total	95	95	**190**	**100%**
FIGURATIVE	5	5	10	5%**

*only literal instances ** divided by 200 (including figurative instances)

Seven instances of the PREC2 word order and five of its sister constructions that are coded as PRESENT furthermore carry progressive aspect. This means that approximately half of the utterances are formed in habitual present tense. Scrutinizing the data by word order, the distribution is somewhat even, although the PREC2 order contains more utterances in PAST tense and the PREC1 order contains more utterances coded as EXISTENTIAL/GENERIC. There are no FUTURE tokens within the PREC2 order, due to word order rules. All but one of the seven concordance lines in the FUTURE belong to the LITERARY genre. Furthermore, four have some religious text as their source and are prophetic in nature.

3.4.1.3 The SPATIAL modifier variable

The distribution and frequencies of the SPATIAL modifier variable within the [*Hagel*+*fallen*] construction are given in Table 80. Even though we already saw that the [ARG+*fallen*] construction with other precipitate types, too, has a proclivity to occur with a variety of SPATIAL modifiers, the [*Hagel*+*fallen*] construction takes this tendency to the next level.The present construction not only encompasses a large number of GOALS, SOURCES and/or PATHS, but also many combinations of SPATIAL modifiers, such as SOURCE+GOAL or LOCATION+PATH+GOAL.

Table 80. *Raw and relative frequencies of the* SPATIAL *modifier variable within the ARG+*fallen*-Cx and the precipitate* Hagel *'hail'*

Word order/	**PREC1**	**PREC2**	**Total**	**Relative frequency**
SPATIAL modifier	*Hagel+fallen*	*fallen+Hagel*		/190
LOCATION	28	35	**63**	**33%**
GOAL	27	18	**45**	**24%**
SOURCE	22	24	**46**	**24%**
PATH	8	4	**12**	**6%**
ZERO	26	24	**50**	**26%**
Total*	**111**	**105**	**216**	**113%**
COMBINATIONS	15	11	26	14%
FIGURATIVE	5	5	10	5%

*The Total includes combinations that are counted individually

The marking of LOCATION is overall the most frequent type of SPATIAL modification, extending to one third of the returns (N=63; 33%). However, only two-thirds of these directly modify the ME (N=44), while one-third appear in the wider context (N=19). About one quarter of the data does not contain any SPATIAL modification (ZERO, N=50; 26%) at all. GOALS and SOURCES are equally frequent and occur in almost one quarter of the returns each, while PATHS have the lowest incidence (N=12; 6%) which is still a noticeable amount compared to the overall low frequency of PATHS across all of the surveyed precipitates. There is a fairly high number of COMBINATIONS (N=26; 14%), SPATIAL modifiers in combinations that are particularly frequent are SOURCES (N=17) and GOALS (N=14) combined (N=8) as well as SOURCES combined with LOCATIONS (N=8). PATHS tend to be combined with either LOCATIONS or GOALS but not with SOURCES (N=1, where the SOURCE furthermore occurs in the wider context). There are two concordance lines where each three SPATIAL modifiers are combined, i.e., "L+S+G" and "S+G+P", as given in (207), where the individual SPATIAL modifiers are indicated by superscript. These are both from the LITERARY genre; the first appears to be prose and the second seems to be from a religious text. With respect to word order differences, the SPATIAL modifiers are fairly homogeneously distributed between the two word orders. The only noticeable difference is between GOALS where the PREC1 order contains 27 tokens and the PREC2 order only 18 tokens.

(206) *Dennoch kam es zu Bildung **einer Superzelle**, **die** am Alpenrand **Hagel** bis 6,5cm **fallen ließ**.* (nature-motive.de)
'It still came to the formation of **a super cell which let hail** of up to 6.5cm **fall**.'

(207) *Da **fiel** unverzüglich **ein schwerer Hagel von oben**^{SOURCE} **auf sie**^{GOAL} **herab**^{PATH} und sie lagen alle tot zu meinen Füßen*. (sos-lehre-zum-leben.de 'sos-teachings-for-life')
'Immediately, **a severe hail fell down**^{PATH} **onto them**^{GOAL} **from above**^{SOURCE} and they all lay dead at my feet.'

Examples of SOURCES are mostly *aus dem/vom Himmel* 'out of/from the skies' (N=18) and and variations of *aus den Wolken*/(*Super*)*zellen* 'out of the clouds/(super)cells', as in (206), where the super cell, furthermore, takes on an agentive role. GOALS are more varied —again, the sources where precipitation comes from are limited, while the goals that precipitation can fall on are basically limitless – examples are the more general *zu Boden* 'to the ground', *auf die Erde/zur Erde* '[on]to the earth' (N=20), *auf den Asphalt/ Straße '*onto the asphalt/streets'*, auf die Menschen/Tiere/Pflanzen* 'onto men/people/animals/plants', as in (205) above. More specific GOALS are such as *aufs Schiff* 'onto the ship', *auf ein bestimmtes Einzugsgebiet* 'onto a specific catchment area', *auf seine Schultern* 'onto his shoulders', *auf den Kopf* 'on the head', and *darauf* 'on it'. Instances of PATHS are *auf sie herab* 'down onto them' (GOAL+PATH, as in (207)), *nieder/herab* 'down', *an den Fenstern vorbeigefegt* 'swept past the windows', or *durch die finstre Luft hernieder* 'down through the gloomy air'. Six concordance lines contain complex paths containing both SOURCE and GOAL (*herab vom Himmel auf die Menschen* 'down from heaven upon the people'. German *Himmel* translates to both 'heaven(s)' and 'sky'. The concordance lines sourced from religious texts probably translate best as 'heaven'—as in "God's abode"—and the concordance lines sourced from the meteorological genre perhaps better translate as 'sky'.) Regarding the semantics of the tokens coded as LOCATION, other than the above mentioned GOALS, SOURCES, and PATHS, there are "specific places", such as *in Burgdorf*, *in ganz Ägypten* 'in all of Egypt', *am Alpenrand* 'at the edge of the Alps', or *in manchen Stadtteilen* 'in some districts', "combinations", and "other"."Combinations" contain, for example, deictics and cardinal directions, such as *örtlich … im Nordosten Deutschlands, in der Westhälfte und in Bayern* 'locally … in the north-east of Germany, in the western half and in Bavaria'. The group "other" is a catch-all for a couple of deictics, e.g., *vielerorts* 'many places' and *dort* 'there', but also items like altitude (*über 800 Meter* 'above 800m'). Most modifiers of PLACE are nominals. Lastly, LOCATION is the most frequent modifier of either type, SPATIAL and NON-SPATIAL, to appear by itself (N=29), followed by QUALITATIVE modifiers (N=13). More than half of the standalone LOCATIVES (N=15) pertain to utterances explaining how weather works.

3.4.1.4 The NON-SPATIAL modifier variable

The distribution and frequencies of NON-SPATIAL modifiers within the [*Hagel+fallen*] construction are given in Table 81. Interestingly, it is not modifiers of TIME with the highest incidence, but modifiers of QUALITY (or MANNER), which are present in more than half of the data (N=107; 56%).

Modifiers of TIME have 21 percentage points less (N=67; 35%). Modifiers of STANCE appear in more than one quarter of the returns (N=53; 28%) and the lowest incidence comes from modifiers of QUANTITY (N=22; 11.5%). Three percent of the data (N=6) do not contain any modifier whatsoever, but more than two-thirds of the returns (N=131; 69%) contain combinations of at least two modifiers of different types, SPATIAL or NON-SPATIAL.

Table 81. *Raw and relative frequencies of the GENRE variable for the ARG+*fallen*-Cx and the precipitate* Hagel *'hail'*

Word order/ **NON-SPATIAL modifier**	**PREC1** *Hagel+fallen*	**PREC2** *fallen+Hagel*	**Total**	**Relative frequency** /190
TEMPORAL	24	42	**66**	**35%**
QUANTITY	10	12	**22**	**11.5%**
QUALITY	48	59	**107**	**56%**
STANCE	22	33	**55**	**29%**
Total	**104**	**146**	**250**	**131.5%**
COMBINATIONS	54	77	131	69%
ZERO	5	1	6	3%
FIGURATIVE	5	5	10	5%
SPATIAL	65	62	127	67%

QUALITY is the most frequent standalone NON-SPATIAL modifier (N=13) and this type also appears most frequently in combination with a LOCATION (N=35). There are an additional three figurative concordance lines modified by a LOCATION and a QUALITATIVE modifier. The nature of each type of NON-SPATIAL modifier is elaborated on in the following sections starting with the modifiers of TIME.

Modifiers of TIME. Temporal modifiers occur in more than one-third of the returned concordance lines and make up about one-fifth of all modifiers (18%). Semantically, they can be grouped into six categories: "sequential", "deictics", "timespan", "date/time", "frequency", and "other". There is not a single category that contains considerably more tokens than the others. "Sequential" and "deictic" items are evenly distributed (N=13 each). "Sequential" tokens recur frequently, such as *später* 'later', *anfangs … später* 'initially … later', *erst … dann* 'first … then', but these are also unique tokens such as *zuletzt* 'finally' and *dazwischen* 'inbetween' as well as the more idiosyncratic *nach seinem Tod* 'after his death'. Since hail seldom occurs by itself, it makes sense to observe a succession of MEs, or different precipitates, and report on them in order, especially in an ME that becomes successively more severe. The "deictic" tokens are all unique, e.g., *an einem Tag* 'one day', *irgendwann* 'sometime', and *um diese Zeit* 'at this time', except for *bald* 'soon' (N=2). The category "duration" (N=11) encompasses both the exact time given in numbers, as in *von 15:35–15:45 Uhr* 'from 3:35pm–3:45pm', as well as lexicalized, less

precise durations such as *innerhalb weniger Minuten* 'within a few minutes' and *innerhalb kurzer Zeit* 'within a short amount of time'. There are four temporal PPs introduced by *innerhalb* 'within' with an added, relatively short duration. In addition to the exact duration, there is a category containing "date/time" items, such as *am 14/15.* 'on the 14th/15th, *1857, am 28. Juli abends 6 Uhr* '1857, on July 28th in the evening 6 o'clock'. Items of "frequency" are less common (N=8): *oft*/*öfter* 'more often/oftener/with increasing frequency' (N=3), *selten* 'rarely' (N=2), and each one token of *niemals* 'never', *immer wieder* 'again and again', and *manchmal* 'sometimes'. Examples of "other", are idiosyncratic items such as *bey* [sic] *den verdoppelten Schlägen* 'at the doubled strokes (of a clock)', *bei Nordsturms Sausen* 'at north storm's swooshing', *erst als ich wieder im Funkhaus angekommen war* 'only when I got back to the radio station', which, with more context, could also be sequential. Otherwise frequent temporal modifiers that are not prevalent in the *Hagel+fallen* returns and are listed as "other" are "day of the week" and/or "time of day" (*am Pfingstmontag Nachmittag* 'on the afternoon of Whitmonday', N=1), "months" (*erst im September* 'not until September', N=1) and "seasons" (*im Winter* 'in winter', N=1).

Modifiers of QUANTITY. With only 11.5% of the returns for *Hagel* 'hail' (N=22), and 6% of all modifiers, QUANTITY is the least frequent. The tokens can be grouped into three categories: "negation" (or "zero quantity"), "large quantities", and "small quantities"/ "other". Roughly half of the quantitative modifiers (N=10) contain negation *kein* 'no' *weder Regen noch Hagel* 'neither rain nor hail' and *nicht* 'not', i.e., the quantity is zero. Roughly half of the QUANTITATIVE modifiers (N=10) lexicalize large quantities by means of nominal modifiers, such as *eine große Menge* 'a big amount' and *Riesenmengen* 'huge amounts'. Other large amounts are described by a result (*so viel Hagel, dass* 'so much hail that…') as in *so viel Hagel, dass er nur noch mit Radladern weggeschafft werden konnte* 'so much hail that it could only be removed with wheel loaders' or *so viel Hagel, dass man im Stadtzentrum meinen konnte, es läge Schnee* 'so much hail, that in the city center you could think there was snow'. Regarding the smaller quantities, these are lexicalized with adverbs *etwas* 'some/a little' *and nur wenig* 'only little'. Additionally, there is one hard number: *10% … fällt dort als Niederschlag wie Regen, Nebel, Hagel, oder Schnee*. '10% …. falls there as precipitation such as rain, fog, hail, or snow'. Obviously, the latter refers to general precipitation, which may fall in different aggregates with hail being only one type. The collocations *fallen+als* 'fall as' and/or *fallen+mit* 'fall with' are discussed in the following section.

Modifiers of QUALITY. More than half of the data contain modifiers of QUALITY. The tokens can be divided into six categories, namely: "size in cm/numbers", "large size", "density", "mixed precipitate", "manner/combinations", and "attributes/other". The size of the hailstone is most frequently lexicalized in as adjectives like *groß*/*sehr groß* 'big/very big' and *klein* 'small' where lexicalizations of large size hail prevail (N=22). The most frequent type is adjective *groß* 'large' and variations thereof, often in combination with the size of the hail stone in centimeters, e.g., *großer Hagel über 3cm Durchmesser* 'large hail over 3cm in diameter'. The size of the hailstones in centimeters alone occurs in six instances. Moreover, the size is often complemented by a comparison: *so groß wie Hühnereier* 'as big as chicken eggs', *wie Pfundstücke/Zentnersteine* 'like pound pieces/hundredweights' and *in der Größe von Taubeneiern und einer solchen Dichte* 'the size of pigeon eggs and the same density'. The last example also refers to density, which warrants its own category, where the manner of falling is denoted by variations of adjective and adverb *dicht* 'dense(ly)': *dichter Hagel fiel auf uns herab* 'dense hail fell down on us', *Hagel fiel so dicht, daß sie das unmittelbar neben sich liegende Boot des Kameraden nicht sahen* 'hail fell so densely that they could not see their comrade's boat lying right next to them'. Here, too, the size in centimeters may be added *(der Hagel fiel) mit ca. 3 cm Durchmesser kleiner aber dafür viel dichter* 'the hail fell with ca. 3cm in diameter smaller but much more densely'. Another seven tokens refer to the size in centimeters, one compares the diameter to that of a 2-Euro coin. The category "mixed precipitate" (N=7) contains tokens either from religious texts (the hail is described as *mit Blut* and/or *Feuer vermischt* 'mixed with blood and/or fire') or, where the concordance is coded as either genre type WEATHER TALK or WEATHER REPORT, the hail is mundanely mixed with rain. Take note that *Regen* 'rain' is not necessarily the main ME as in *anfangs war der Hagel mit schwerem Regen vermischt später fiel aber nur Hagel* 'initially the hail was mixed with heavy rain but later solely hail fell'.

Qualitative attributes are furthermore denoted as *eisig* 'icy', *trocken*'dry', *fein* 'fine', *dick* 'thick', *schwer* 'heavy', *weich* 'soft', *gelb* 'yellow' (sourced from what presumably is a fictional story), and *brennend* 'burning'. A couple of adjectives entail a more forceful negative semantic prosody: the precipitate is characterized as *zerschmetternd* 'crushing', *furchtbar* or *schrecklich* both 'awful/terrible' and *große Verheerungen anrichtend* 'causing great devastation'. These stance-imbued adjectives usually pertain to concordance lines from the religious genre. Lastly, the manner in which hail falls (other than *dicht* 'densely') is described as *anhaltend* 'persistend', *langsam* 'slow', and *reichhaltig* 'ample'. Where QUALITY is the only modifier, hail is almost

exclusively described as *groß* 'big' or *dicht* 'dense'. Except for one token, all standalone QUALITATIVE modifiers belong to the PREC1 word order.

Modifiers of STANCE. Almost one-third of the corpus returns for *Hagel* 'hail' expressions are imbued with stance. Surprisingly, there is one instance of positive semantic prosody with the speaker lamenting not having seen hail even once all through the year. Four utterances are mirative and the remainder, the bulk of STANCE expressions, unsurprisingly carry negative semantic prosody. Many of the negative concordance lines are LITERARY and from religious sources, as in (208), where hail is generally sent as some sort of punishment and in the worst case does not only destroy crops but kills people and animals. Of course, hail is the seventh plague in the *Book of Exodus*. Concordance lines pertaining to the WEATHER TALK genre, such as news outlets, generally discuss the damage caused by hail or "private" speakers expose the damage done to their plants and gardens, as in (209), as well as bodily harm experienced by speakers caught in a hailstorm during outdoor activities, as in (210).

(208) *… alles Vieh, die auf dem Felde gefunden und nicht ins Haus aufgenommen werden,* ***auf die fällt der Hagel herab, und sie werden sterben***. (bibelkommentare.de 'Bible commentaries')
'… **the hail will fall down onto** every animal that has not been brought in and is still out in the field, **and they will die**.'

(209) *Besonders im Raum Litschau* ***fiel*** *eine große Menge* ***Hagel, wodurch viele Schäden*** *an Pflanzen* ***angerichtet wurden***. (meinbezirk.at 'my district')
'A large amount of **hail fell**, especially in the Litschau area, **which caused a lot of damage** to plants.'

(210) *Aber wenn kurz nach dem Start ohne Ankündigung ganz feiner* ***Hagel*** *vom Himmel* ***zu fallen beginnt und in meinem Gesicht schmerzt****, dann verlassen mich die Freuden des Tages.*
(laufhannes.de 'run Hannes')
'But when, shortly after the start, without warning, a very fine **hail begins to fall** from the sky **and hurts my face**, then the joys of the day desert me.'

Stance is clearly communicated through many of the qualitative modifiers mentioned above, e.g., *furchtbar/schrecklich* 'terrible', *zerschmetternd* 'crushing', or multiword modifiers such as *große Verheerungen anrichtend* 'causing great devastation', but it can also be expressed through descriptions of the damage that the falling hail caused, as in (209) or more implicitly through lexial items like *betroffen* 'affected'.

3.4.1.5 The co-ME (co-occurring meteorological event) variable: MIT Hagel *'with hail' and* ALS Hagel *'as hail' /* in Form von *'in the shape of'*

There are 27 concordance lines, where some precipitate or other figure falls *als Hagel* 'as hail' and seven concordance lines, where some precipitate falls *mit Hagel* 'with hail'. Mostly the source substance for the hail fall is given as generic *Niederschlag* 'precipitation'—which occurs predominantly in singular (N=7) and less so in plural (N=1)—but also of *Wasser* 'water' (N=5), *Wassertropfen* 'water drops' (N=4), and each one token of *Wasserpartikel* 'water particle', *Schluck Wasser* 'a drink of water', *Wasserdampf* 'water vapour', and *Nebel und Wasserdampf* 'fog and water vapour'. Additionally, there is one token each of *Regen* 'rain', *Regenschauer* 'rain showers', *Schnee* 'snow', *Tröpfchen* 'droplets', *Eisstücke* 'pieces of ice', and *Eiskörper* 'bodies of ice'. More often than not, the figure falling is followed by a list of possible byproducts in different aggregates (*Wasser fällt als Regen, Hagel, oder Schnee* 'water falls as rain, hail, or snow'), but there are also a couple of concordance lines, where *Hagel* 'hail' is the sole aggregate (*Schnee fällt als Hagel* 'snow falls as hail').

Regarding the interaction with other variables, the majority of concordance lines are EXISTENTIAL/GENERIC (N=22) —unsurprisingly, 17 of belong to the METEOROLOGICAL genre— and some are in PAST tense (N=5). Less frequent genre types are WEATHER TALK (N=5), LITERARY (N=3) and WEATHER REPORTS (N=2). One of the WEATHER TALK examples is given in (211) where the speaker lists four precipitates and, after the target construction with *fallen* 'fall', additionally lexicalizes MEs by using a *es*+PRED-Cx as well as some other ARG-Cx with light verb *kommen* 'come'. Take note that the first ME here (*da gibt es nicht nur* Regen 'not just rain exists') is a "true" existential *geben* 'give' construction in that it refers to the existence of precipitates.

(211) *Aber da gibt es nicht nur Regen – Gottes Schöpfung ist nicht langweilig, sondern Gott hat für unseren Schluck Wasser eine wundervolle Vielfalt parat,* ***wie er aus der Wolke fallen kann****:* ***als*** *Schnee,* ***Hagel****, Graupel oder Nebel. Da nieselt es oder da schüttet es; da kommt ein Schauer oder ein Dauerregen*. (predigtkasten.de 'sermon box')
'But there is not only rain - God's creation is not boring, but God has a wonderful variety at the ready for our drink of water, of **how it can fall from the cloud: as** snow, **hail**, graupel (NB: soft hail) or fog. It's drizzling or it's pouring; there comes a shower or a steady rain.'

(212) *Furchtbarer Donner grollte, Blitze zuckten,* ***es fiel ein starker Regen mit Hagel****. Die Körner waren so groß, daß es den Leuten vorkam, als würden sie mit Stöcken auf die Beine geschlagen.* (kondor.de)
'Terrible thunder rumbled, lightning flashed, **it fell a heavy rain with hail**. The grains were so large that people felt as if they were being hit on the legs with sticks.'

The *Hagel* 'hail' precipitate does also not take center stage where different precipitates fall concomitantly: *mit Hagel* 'with hail' (N=8) or *vermischt mit Hagel* 'mixed with hail'. *Mit Hagel* returns tend to be the co-ME of *Regen* 'rain', as in (212), or *Wassermassen* 'masses of water', but not *Gewitter* 'thunderstorms', although hail forms inside thunderstorm updrafts. However, each six concordance lines mention thunderstorms in the left and right context. In total, there are only 82 corpus returns for the [*Hagel*+*fallen*] construction where the focus is on *Hagel* alone and no other co-MEs or types of precipitation are mentioned.

3.4.2 The argument construction [Hagel+geben] *'hail'+'give'*

This section surveys the ARG-Cx [*Hagel*+*geben*] ('there is hail', lit. 'hail'+'give'). The queries in the *deTenTen13* were carried out using the regular expressions in (xxvii) and (xxviii) to account for the different word orders. The search returned 979 concordance lines in total. However, the comparatively high number of returns is misleading since a total of 235 concordance lines (87 for the PREC1 order and 148 for the PREC2 order) had to be removed as false returns.

(xxvii) [word="Hagel"] [ ] {0,8} [lemma="geben"] within <s/> (N=389)
(xxviii) [lemma="geben"] [ ] {0,8} [word="Hagel"] within <s/> (N=590)

3.4.2.1 USAGE (LITERAL VS FIGURATIVE) AND GENRE VARIABLES

The [ARG+*geben*] construction with precipitate *Hagel* 'hail' overwhelmingly yields literal results (N=194, cf. Table 82). There are six concordance lines where the construction is used figuratively. In all six utterances the noun *Hagel* 'hail' is complemented by another mostly negative noun, e.g., *ein Hagel an Wurfgeschossen* 'a hail of missiles', *ein Hagel von Zeitstrafen* 'a hail of time penalties', *ein Hagel von Schlägen*/*Pfeilen*/*Protesten*/*Anrufen* 'a hail of punches/arrows/protests/calls'.

Table 82. *Frequencies of literal and figurative USE for the ARG+*geben-Cx *and the precipitate* Hagel *'hail'*

Word order/ USAGE	**PREC1** *Hagel+geben*	**PREC2** *geben+Hagel*	**Total**
Literal	97	97	**194**
Figurative	3	3	**6**
Total	**100**	**100**	**200**

The absolute and relative frequencies of the [ARG+*geben*] construction and precipitate *Hagel* 'hail' and their distribution by GENRE are given in Table 83. This constxsruction is predominantly attracted by genre type WEATHER TALK with (N=101; 52%). WEATHER REPORTS are fairly frequent and come second with about one-third of the returns (N=73; 38%). Third most frequent, by a wide margin, are concordance lines from the METEOROLOGICAL genre (N=13; 7%) followed by LITERARY utterances (N=6; 3%). There is a single utterance coded as OTHER. The remainder are FIGURATIVE MEs.

Table 83. *Raw and relative frequencies of the GENRE variable for the ARG+*geben*-Cx and the precipitate* Hagel *'hail'*

Word order/	**PREC1**	**PREC2**	**Total**	**Relative frequency**
GENRE	*Hagel+geben*	*geben+Hagel*		/194*
WEATHER TALK	53	48	**101**	**52%**
WEATHER REPORT	31	42	**73**	**38%**
METEOROLOGY	8	5	**13**	**7%**
LITERATURE	4	2	**6**	**3%**
OTHER	1	0	**1**	**0.5%**
Total	**97**	**97**	**194**	**100.5%**
FIGURATIVE	3	3	6	divided by 200: 3%

*literal instances

By word order, the low frequency genres LITERARY and METEOROLOGY are distributed somewhat evenly. While the PREC2 order is fairly homogeneously distributed between WEATHER TALK and WEATHER REPORTS, its sister construction contains roughly one and a half times as many constructions belonging to WEATHER TALK as opposed to WEATHER REPORTS.

3.4.2.2 The TAM (TENSE, ASPECT, MOOD) Variable

The frequencies of the [*Hagel+geben*] construction and their distribution by TAM level are given in Table 84. The construction is predominantly used in PAST tense (N=137; 70.5%). Trailing far behind with the second highest incidence come instances with "future intent" (N=26; 14%); most of these are coded present_FUTUR (N=21; 11%) and only a handful are coded for the periphrastic FUTURE (N=4) and IRREALIS (N=1). EXISTENTIAL/GENERIC use has a similar incidence as future intent with only two percentage points less (N=23; 12%). Lastly, there are only eight instances where the utterance is inflected for PRESENT tense. Only one of these instances contains progressive marking.

Table 84. *Raw and relative frequencies of the* TAM *variable for the ARG+*geben*-Cx and the precipitate* Hagel *'hail'*

Word order/ TENSE	**PREC1** *Hagel+geben*	**PREC2** *geben+Hagel*	**Total**	**Relative frequency** /194
PAST	53	84	**137**	**70.5%**
PRESENT	4	4	**8**	**4%**
present_FUTURE	20	1	**21**	**11%**
FUTURE	4	0	**4**	**2%**
IRREALIS	1	0	**1**	**0.5%**
EXIS _ GEN	15	8	**23**	**12%**
Total	**97**	**97**	**194**	**100%**
FIGURATIVE	3	3	6	3%

Looking at the construction by word order, there are two findings that stand out. First, regarding the use of present tense indicating a future event, 20 of the 21 utterances are found in the PREC1 data, mostly due to the fact that they contain a modal verb, as in (213). The modal verbs are *kann* 'can' (N=8), *könnte* 'could' (N=2), *soll* 'should/ought' (N=5), and *sollte* 'was supposed to' (N=1). Unsurprisingly, 10 of these concordance lines belong to genre level WEATHER REPORT, which means they are predictions that express a possibility. While the remaining five concordance lines still refer to weather forecasts, they are more informal in nature. This is obvious from the use of embedded discourse markers *ja* (lit. 'yes') and *auch* 'also/too' in example (214).

(213) *Außerdem **kann es Hagel** und schwere Sturmböen **geben**.* (eichwaelder.de)
'Additionally, **it can give hail** and heavy storm.'

(214) *Heute **soll es** ja auch Unwetter mit **Hagel geben**.* (agrar.de/pferde/forum 'agrarian/horse/forum')[94]
'**It's supposed** to **give** thunderstorms with **hail** today.'

Take note that in the example above, *Hagel* 'hail' is a co-ME of the forecasted storm (*Unwetter mit Hagel* 'thunderstorm with hail') and thus is not necessarily the focus of the ME (due to the simple fact that hail tends to be a byproduct of storms and not the other way round). Lastly, the four concordance lines that do not contain any modal verb are negated in some way and are furthermore conditional in nature, as in (215), with two of them being *wenn* 'if' clauses, as in (216).

[94] The particles are epistemic in nature: *ja* generally indicates that both speaker and hearer already know the proposition (cf. Karagjosova 2003), even if the speaker does not make such assumptions, *auch* 'also' refers backwards.

(215) *Naja, solange **es** keinen **Hagel gibt**…* (marathon4you.de)
'Well, as long as **it gives** no **hail**…'

(216) "***Wenn es** in den nächsten Wochen keine Unwetter und **keinen Hagel gibt**, wird es noch reichlich Kirschen geben", versicherte (…) Werner Kipp (…).* (wetterauer-zeitung.de – 'Wetterauer-newspaper')
'**If it gives** no thunderstorms and **no hail** in the coming weeks, there will be plenty of cherries (lit. 'it is going to give plenty of cherries')" assures Werner Kipp.'

Example (216) clearly points to the future by referring to 'the coming weeks' (*in den nächsten Wochen*) while the verb is in present tense. In example (215), there is no such indicator of a future event, but it is obvious that the speaker is not referring to an ongoing weather event. (The source website indicates that the speaker is referring to an upcoming race/a marathon.)

The second finding pertains to the PAST tense results. The tokens in the PREC2 order (N=84) add up to 43% of the returns, while the PREC1 order (N=53) only adds up to 27% of the returns. Furthermore, out of the 84 PAST tokens, only one is in past perfect; the other 83 are in preterite. The utterances are somewhat evenly split between genre types WEATHER TALK and WEATHER REPORT; only one PAST tense token belongs to the LITERARY genre. Obviously, the PREC2 word order contains hardly any "future intent" tokens, which appeared to be a good indicator of genre type weather reports—or rather weather forecasts—which is still valid for the PREC1 order. The concordance lines coded as PAST and WEATHER REPORT in the PREC2 order are indeed "reports" on a past weather event from news outlets and websites dedicated to monitoring the weather (e.g., *wettergefahren-fruehwarnung.de* 'weatherdangers-earlywarning'). Furthermore, more than half of the WEATHER REPORTS hail from the same website, i.e., *sturmwetter.de* ('stormweather'). In contrast, the PREC1 order returns are much more varied with 30 tokens in preterite, 20 tokens in *Perfekt*, where a single instance is furthermore progressive, and, lastly, four utterances are in past perfect. Ten of the returns also come from website sturmwetter.de, but these are not all in preterite, but three are in *Perfekt* tense and one is in past perfect. Otherwise, the low frequencies of FUTURE tense, IRREALIS, and PRESENT tense—especially lack of progressive aspect—stand out.

3.4.2.3 The SPATIAL modifier variable

The frequencies of the [*Hagel+geben*] construction and their distribution by SPATIAL modification are given in Table 85. Regarding the aggregate frequencies, this construction resembles both the

[*es+hageln*] construction and the [*Hagel+fallen*] construction in that LOCATION is not the most frequent modifier. The returns of these construction basically contain either a LOCATION (N=98; 51%) or no SPATIAL modifiers whatsoever (N=94; 48%). The LOCATIONS predominantly directly modify the ME, only in eight utterance is the LOCATION mentioned in the wider context. here are only two minor SPATIAL modifiers, a GOAL and a PATH, which each occur in combination with a LOCATION. The GOAL is *auf unser Haupt* 'onto our head' and the PATH is *Hagel, die über das Stadtgebiet München hinweggezogen sind* 'hail that passed over the city region of Munich'. There are zero utterances where a SOURCE, GOAL, or TRAJECTORY stand alone or where a SOURCE is combined with a LOCATION. The semantic nature of the LOCATIONS is surveyed below.

Table 85. *Raw and relative frequencies of the SPATIAL modifier variable for the ARG+*geben*-Cx and the precipitate* Hagel *'hail'*

Word order / SPATIAL modifier	**PREC1** *Hagel+geben*	**PREC2** *geben+Hagel*	**Total**	**Relative frequency** /194
LOCATION	45	53	**98**	**51%**
GOAL	1	0	**1**	**0.5%**
SOURCE	0	0	**0**	**0%**
PATH	0	1	**1**	**0.5%**
ZERO	51	43	**94**	**48%**
Total	**97**	**97**	**194**	**100%**
COMBINATIONS	1 (L+G)	1 (L+T)	2	1%
FIGURATIVE	3	3	6	3%

The LOCATIVES can be arranged into four groups: mostly the tokens are "location", as in specific places, and "deictics", then there are "combinations", which chiefly contain some cardinal direction as one of the combined modifiers, and lastly "other". The most frequent category is "location" (N=34), which are all unique types except for *in Deutschland* 'in Germany' (N=2). Other than pinpointing cities, countries, and regions (e.g., *in Oberbayern* 'in Upper Bavaria'), for the concordance lines coded as METEOROLOGY, the location of the precipitate is denoted as being inside the ME itself, i.e., *in manchem Gewitter* 'in many a thunderstorm' or *kurz hinter der Aufwindzone* 'just behind the updraft zone'. The second most frequent category, "deictics", (N=32), mostly contains pointers like *hier* 'here' (N=8), *bei uns* 'here' (N=3) and *dort* 'there' (N=5). Another five tokens specify the place denoted by *hier* 'here', as in *hier nahe Saarbrücken* 'here close to Saarbrücken' or *hier oben* 'up here'. Strictly speaking, these are combinations, too. Other tokens are such as *gebietsweise* 'in some areas, *örtlich* 'locally', *mancherorts* 'in some places', and *stellenweise* 'in spots', i.e., syntactically, this category involves mostly adverbs. The category "combinations" (N=15) contains at least two items from two different categories, e.g.,

im Landkreis Rosenheim, dort … gebietsweise 'in the district of Rosenheim, there ... regionally'. This example is one of only three concordance lines, which do not include a cardinal direction. The other 12 tokens all contain a cardinal direction, such as *vom südlichen Rheinland-Pfalz bis nach Sachsen hinein* …. 'from southern Rhineland-Palatinate into Saxony...' or *am südlichen Starnberger See* 'on the southern Lake Starnberg'. The catch-all category "other" encompasses 10 miscellaneous tokens, such as *bei Ilona* 'at Ilona's' and *in den Heimatgebieten dieser Tiere* 'in the habitats of these animals'.

3.4.2.4 The NON-SPATIAL modifier variable

The distribution and frequencies of the NON-SPATIAL modifier variable within the [*Hagel+geben*] construction are given in Table 86. SPATIAL modifiers are given in the last row and count towards rows ZERO and COMBINATIONS. The most frequent modifier is TEMPORAL (N=111; 57%), just as it is for the [*es+hageln*] construction. The second most frequent are modifiers of QUALITY/MANNER (N=77; 40%) followed closely by STANCE (N=64; 33%), which is present in one-third of the data. Obviously, these two modifiers often go hand-in-hand. In line with the results of the other constructions, the modifier with the lowest frequency is QUANTITATIVE in nature, here with 24 tokens in total (12%). Ten concordance lines (5%) contain no modifier whatsoever (ZERO), SPATIAL or NON-SPATIAL.

Table 86. *Raw and relative frequencies of NON-SPATIAL modifiers for the ARG+*geben*-Cx and the precipitate* Hagel *'hail'*

Word order/ NON-SPATIAL modifier	**PREC1** *Hagel+geben*	**PREC2** *geben+Hagel*	**Total**	**Relative frequency** /194
TEMPORAL	56	55	**111**	**57%**
QUANTITY	14	10	**24**	**12%**
QUALITY	39	38	**77**	**40%**
STANCE	33	31	**64**	**33%**
Total	**142**	**134**	**267**	**142%**
COMBINATIONS	63	73	136	70%
ZERO	5	5	10	5%
FIGURATIVE	3	3	6	3%
SPATIAL	40	51	91	47%

The nature of each modifier is elaborated on individually in the following subsections starting with the most frequent, modifiers of TIME.

Modifiers of TIME. TEMPORAL modifiers occur in more than half of the data, which makes them the most frequent NON-SPATIAL modifier. The distribution between the two word orders is

essentially even. In addition to being the most frequent modifier, TEMPORALITY also shows the most diversity with regard to the semantic categories the modifiers fall into as well as the different types. The first of the eight categories, "deictics", is the largest category (N=25) and contains items such as *heute* 'today', *diesmal* 'this time', and *in den letzten Tagen* 'in the last days', where *heute* 'today' is the most frequent type (N=5). The second highest incidence is provided by items of "frequency" (N=18), e.g., *immer wieder* 'again and again', *einen Tag* 'one day', where no type occurs more than twice. Moreover, TEMPORAL modifiers can be grouped as "time of day" (*mittags* 'at midday', *in der Nacht* 'at night'), "day of the week" (*am Sonntag* 'on Sunday', *am Wochenende* 'on the weekend'), "months and seasons" (*im August* 'in August', *im Herbst* 'in autumn'), and "sequence" (*anfangs* 'initially', *kurz darauf* 'shortly after', *kurze Zeit später* 'a short time later'). There is a fairly large group of more complex modifiers (N=16), which combine two or more items from different groups, such as *am Nachmittag des 2. August 2009 zwischen 15:00 Uhr und 15:30 Uhr* 'on the afternoon of August 2, 2009 between 3:00 p.m. and 3:30 p.m' or *am 1. Juli 2011 tagsüber* 'on July 1, during the day'. "Other" items can be something like the year (e.g., *1873*) but more often are idiosyncratic: *nachdem ich den Salat auspflanzte* 'after I planted the lettuce', which is also sequential, *genau zum City Walk* 'right during the City Walk', *während der Inselrallye* 'during the island ralley', *während ich dort war* 'while I was there', *genau in dem Moment, als Helen und ich auf das Seil hoch mussten* 'right at the moment when Helen and I had to go up the rope'.

Other than semantically, the TEMPORAL modifiers can also be grouped according to their grammatical category: nouns (e.g., *gestern Abend* 'yesternight', *Ostersonntag* 'Easter Sunday'), or noun phrases (*am Ende der Aufführung am Sonntag Abend* 'at the end of the performance Sunday night'), adjectives (*früh* 'early', *im nächsten Moment* 'in the next moment'), or adverbs (e.g., *heute* 'today'). Overall, the modifiers did not fit neatly and easily into categories. The number of types is relatively high with no type more frequent than five tokens.

Modifiers of QUANTITY: The majority of the infrequent modifiers of QUANTITY can basically be categorized in two groups: "negation" and "other". There are fifteen tokens of negation (*keinen* 'no', N=12; *nicht* 'not', N=2; and *außer* 'except for', as in (217)).

(217) *Die ganze Wetterpalette (**außer** Schnee und **Hagel**) **gibt** **es** auch in den nächsten Tagen, Sonne und Regenwolken wechseln sich ab.* (langeoognews.de 'Langeoog news')
'In the next days, **it gives** the whole weather palette (**except for** snow and **hail'**), sun and rain clouds alternate.'

The category "other" basically encompasses all quantities that are positive (as opposed to the negated ones). Each is a unique type, such as *ein bisschen* 'a little' and *den meisten* 'the most', *wesentlich weniger* 'considerably less', and the collocational phrase *und das nicht zu knapp!!* 'and how!!', meaning 'in abundance'. There is also one concordance focusing on the product of the ME: *von bis zu einem Meter Höhe* 'of up to one meter height' – which seems a little excessive (if not unbelievable) for a hail event. In line with the modification of the main event (where *hail* is a co-ME), it is also the thunderstorm that is quantified rather than the hail itself as in *zahlreiche Gewitter mit Hagel* 'plenty of thunderstorms with hail' or *einige Unwetter mit…* 'several thunderstorms with…'. The high number of negations for the small amount of quantitative modifiers is furthermore in line with the tendency of the [ARG+*geben*] constructions to interact with negation, as we have seen for other precipitates.

Modifiers of QUALITY. The modification with the second highest incidence, QUALITY and/or MANNER, mostly expresses the size of the hailstones in centimeters (CM, N=30). Another 19 utterances contain the size of the hailstones mostly in the shape of adjectives with the most frequent type being *kleinkörnig* 'fine-grained' (N=8). Another adjective referring to small sized hailstones is *klein* 'small' (N=3). The opposite, i.e., large hailstones, are denoted as *groß*/*sehr groß* 'big/very big' or *größer* 'bigger' (N=4 altogether). With the larger size, speakers also make use of comparisons: the size of hail is described as *tennisballgroß* (N=2) or *in Tennisballgröße*, both '[in] the size of a tennisball', *hühnereigroß* 'chicken-egg-sized' or *in normaler Größe, nicht wie Tennisbälle* 'normal size, not like tennis balls'. That means, apart from the size in numbers, there are two types and ten tokens referring to smaller hailstones and another eight tokens referring to large hailstones. "Medium sized" hailstones tend not to be lexicalized. In addition to attribute size, the QUALITY of hail is described as *dick* 'thick', *mächtig* 'mighty', *furchtbar* 'terrible', *stark* 'strong'/*ziemlich stark* 'pretty strong', *heftig/heftigst* 'severe/severest', *leicht* 'light' and *schwer* 'heavy', *gewaltig* 'enormous', and *astrein* 'great/genuine', informal/slang! or *auch noch Stunden später sichtbar* 'still visible hours later'.

Where hail does not take center stage but is the co-ME of a thunderstorm, it is the thunderstorm that is modified rather than the hail itself, such as *heftige/kräftige/schwere Gewitter/Unwetter* 'violent/severe/heavy thunderstorms' *mit Hagel* 'with hail'. There are 12 instances where *Hagel* 'hail' is the co-ME or byproduct, mostly "added on" by use of preposition *mit* 'with', and where the preceding ME is further modified. These instances are included in the frequency count of QUALITATIVE modifiers as, arguably, the presence of hail is a significant

qualifying factor in making said ME "severe". Still, even if these 12 tokens were removed from the frequency count, modifiers of QUALITY would still be the third most frequent type of modifier overall.

Modifiers of STANCE. Expressions of STANCE occur in one-third of the returns and are almost evenly distributed between the two worder orders. The majority of utterances are unambiguously negative (N=37), such as *furchtbares Wetter*/ *Sche***wetter* 'terrible weather/shi*weather' and *es war noch schlimmer* 'it was even worse'. Another 11 tokens are positive but contain negation, so that they too are negative regarding the ME itself, as in *Gottseidank keinen Hagel* 'thank God no hail'. Nine concordance lines signal surprise, such as mirative *und was für einen* 'and what a [hail]'; six of these contain the particle *sogar* 'even': *neben Regen gab es sogar Hagel* 'besides rain it even gave hail'. The group "other" (N=7) contains, among others, instances of evidentiality; in the following example the sentence additionally ends with an exclamation point, possibly signaling the speaker's unpreparedness of mind: *Ich habe von Beobachtern erfahren, dass…!* 'I learned from observers that…!'. The next subsection discusses the use of precipitate *Hagel* 'hail' as a byproduct or co-ME since it rarely occurs by itself.

3.4.2.5 The co-ME Variable: MIT Hagel *'with hail' and* ALS Hagel *'as hail' /* in Form von *'in the shape of'*

As mentioned earlier, hail forms inside thunderstorm updrafts and is therefore often mentioned as the byproduct or co-ME of some other ME rather than taking center stage. Accordingly, there are 43 concordance lines that contain the PP *mit+Hagel* 'with+hail'. Again, the distribution between word orders is surprisingly even. The most frequent ME to occur as the main event is *Gewitter* 'thunderstorm' (N=27). Furthermore, there are instances of *Unwetter* 'severe weather/thunderstorm' (lit. "unweather", N=6), *Platzregen* 'downpour' (N=3), and *Sturm* 'storm' (N=2). The following "main event MEs" occur each once: *Mesozyklone* 'mesocyclones', *Wolkenbruch* 'cloud burst', *Sturzregen* 'downpour', and *Unwetterwarnung* 'severe weather warning'. Lastly, there is one instance of *Regenschauer* 'rain shower', where the verb *begleiten* 'accompany' appears in addition to preposition *mit* 'with'. Usually, *begleiten* would occur with the preposition *von* 'by', however, the concordance lines stems from an Austrian source, so that the combination might be a regiolect.

Table 87 shows that, as opposed to the [*Hagel*+*fallen*] construction, there is not a single instance of some precipitate or precipitation of *geben als Hagel* 'give as hail' in the coded data, which would result in an ill-formed sentence [**es gibt Regen als Hagel* 'it gives rain as hail').

***Table 87**. Distribution and frequencies of the precipitate* Hagel *'hail' as a co_ME in combination with the preposition* mit *'with' and the conjunction* als *'as' within the returns of the ARG-Cxx*

preposition / conj	*Hagel+fallen*	*Hagel+geben*	Total
mit **'with'**	8	43	**51**
als **'as'**	29	0	**29**
Total	**37**	**43**	**80**

Meanwhile, the [*Hagel*+*fallen*] construction overwhelmingly contains concordance lines with conjunction *als* 'as' but also concordance lines with preposition *mit* 'with', albeit to a much lesser extent. Interestingly, the total of the co-MEs is relatively evenly distributed between the two constructions with a difference of only six tokens (*Hagel*+*fallen* N=37 vs. *Hagel*+*geben* N=43). This paragraph concludes the survey of the [ARG+*geben*] construction in combination with precipitate *Hagel* 'hail'. The next subsection surveys the [*es*+*hageln*] construction, which, in at least one regard, differs profoundly from the *es*+PRED-Cx with other precipitates.

3.4.3 The predicate construction [es+hagelt] *'it+hail'*

This subsection deals with the [*es*+*hageln*] construction. The search queries in (xxix) and (xxx) were carried out to account for both word orders. The search for these two queries returned 22,073 concordance lines in total. Strikingly, the query (xxx) returned almost twice as many results as query (xxix).

(xxix) [word="es"] [] {0,8} [lemma="hageln"] within <s/> (N=7,589)
(xxx) [lemma="hageln"] [] {0,8} [word="es"] within <s/> (N=14,484)

Only one concordance line had to be removed as the utterance did not make sense. However, an unexpected occurrence, discussed in the following section, reduced the amount of "codeable" concordance lines considerably.

3.4.3.1 USAGE: LITERAL VS FIGURATIVE and the variable GENRE

The instant and most striking finding for the [*es+hageln*] construction is the overwhelming figurative use of the verb. The output for the (xxx) search query only contains two instances of literal use. The output for the (xxix) serach query also shows a substantial majority of non-literal usage. Here, 23 instances of the [*es+hageln*] construction are used literally (cf. Table 88). That means, only 25 concordance lines in total (12.5%) are used literally. The following tables focus on concordance lines with literal usage. The preponderance of figurative usages is explored in Chapter 4, Section 4.4. The data were coded for the TAM variable as well as the direct objects that follow the predicate *hageln* 'hail'.

Table 88. *Frequencies of literal and figurative USAGE for the* es*+PRED-Cx and the precipitate* Hagel *'hail'*

Word order /	**PREC1**	**PREC2**	**Total**
Literal	2	23	25
Figurative	98	77	175
Total	**25**	**175**	**200**

The distribution and frequencies of the GENRE variable are given in Table 89. Due to the overwhelming figurative use of the construction (N=175; 87.5%), the types of GENRE differ from those of other precipitates. Most noticeable is the absence of both genre types WEATHER REPORTS and METEOROLOGICAL uses and only one token each that is LITERARY and OTHER. Where the precipitate is used in its literal meaning, it is almost always part of genre type WEATHER TALK (N=23).

Table 89. *Raw and relative frequencies of the GENRE variable in the* es*+PRED-Cx and the precipitate* Hagel *'hail'*

Word order/	**PREC2**	**PREC1**	**Total**	**Relative frequency**
GENRE	*es+hageln*	*hageln+es*		/200
WEATHER TALK	21	2	**23**	**11.5%**
WEATHER REPORT	0	0	**0**	**0%**
METEOROLOGY	0	0	**0**	**0%**
LITERARY	1	0	**1**	**5%**
OTHER	1	0	**1**	**5%**
FIGURATIVE	77	98	**175**	**87.5%**
Total	100	100	**200**	**100%**

The token from the LITERARY source is some sort of retelling of a Greek legend. The token coded as OTHER is an instance where the author discusses *Wetter-Verben* 'weather verbs' and *Nullwertigkeit* 'zero-valency', given in (218) below. It is possible that the sample size is simply

too small to find more variation, such as instances about AVW or HWW. It is also possible that, as it would be difficult to predict hail, there are no or few instances of genre types other than WEATHER TALK. Although some geographical areas apparently are more prone to experiencing thunderstorms, the occurrence of the precipitate seems to be rare and random enough to not warrant any mention. We shall see in the description of the argument constructions whether there is more variety.

3.4.3.2 The TAM (TENSE, ASPECT, MOOD) variable

The distribution and frequencies of the TAM variable for literal meaning by construction are given in Table 90. Due to the low overall frequencies, the relative frequencies are not provided. The TAM levels present_FUTURE, FUTURE, and IRREALIS are subsumed under "future intent". The two results for the PREC1 order are both in preterite. There is more variation in the PREC2 order returns despite the overall low results. This is true for both literal and figurative usage. The figurative (xxx) query results are exclusively either in past or present tense. The PREC2 order results contain at least one return for each of the tenses coded for except for FUTURE. Future tense only occurs with figurative usage and so does IRREALIS (N=10), except for one example that is used in a literal manner.

Table 90. *Raw frequencies of the TAM variable within literal instances of the* es*+PRED-Cx and the precipitate Hagel 'hail'*

Word order / TAM level	**PREC1** *hageln+es*	**PREC2** *es+hageln*	**Total**
PAST	2	10	**12**
PRESENT	0	7	**7**
FUTURE_intent	0	5	**5**
EXIS_GEN	0	1	**1**
Total	**2**	**23**	**25**

Approximately half of the concordance lines (N=10) are coded as PAST. Five of these are additionally coded as *Perfekt*, one is additionally marked for progressive aspect and another two are infinitives in combination with a quasi-auxiliary indicating incipient aspect (i.e., *es begann zu hageln* 'it started to hail'). Seven concordance lines occur in simple present where three are habitual, two furthermore carry progressive aspect, and another two are infinitives in combination with quasi-auxiliaries *beginnen*/*anfangen* 'begin/start'. Four concordance lines are coded as present_FUTURE for the German *Konjunktiv II*, the general subjunctive that (among

other usages) expresses hypothetical situations and conditionals and is formed with the help of auxiliary or modal verbs. There is one concordance coded as EXISTENTIAL/GENERIC where the author discusses the use of weather verbs, claiming that they can occur without any verb complement, as in (218). Although the utterance is about weather verbs rather than an actual weather event, the examples themselves (*es schneit, es regnet, es hagelt*) still refer to actual weather events rather than figurative ones, so that the concordance lines is coded as "literal". Last but not least, there is one utterance in "true" IRREALIS where the speaker wishes that 'it would hail cats outside' (*es würde draußen Katzen hageln*), so that they had an excuse to stay home instead of having to run running errands. This concordance is also a unique example of positive semantic prosody. There are zero instances of FUTURE tense (*es wird hageln* 'it is going to hail) within the literal data.

(218) *Dabei sind die Wetter-Verben so seltene Vögel, kommen sie doch ohne Verb-Ergänzung aus: es schneit, es regnet,* ***es hagelt****.* (irgendlink.de – 'somelink')
'And yet, weather verbs are such rare birds, they manage without verb complements: it snows, it rains, **it hails**.'

Overall, for such a small data set, there is quite some variation with regard to TAM marking, which may, in combination with the genre type WEATHER TALK, hint at the ME being lexicalized more informally and "conceptually oral" rather than written. Regarding the topics of the sources that the utterances come from, approximately half of the literal concordance lines refer to outdoor events such as concerts and sports competitions, two are about language, as in (218), two are from news blogs (fire department and regional news), and the remainder have some kind of personal blog as source material.

3.4.3.3 The SPATIAL modifier variable

The distribution and frequencies of the SPATIAL modifier variable by word order are given in Table 91 below. Of the 23 literal concordance lines, 16 do not encompass any SPATIAL modifier whatsoever (ZERO). Nine concordance lines contain a LOCATION, but only five of the locations directly modify the ME. The remainder occurs in the wider context. There are zero instances of "minor" SPATIAL modifiers GOAL, SOURCE, or PATH, and, correspondingly, zero COMBINATIONS.

Table 91. Frequencies of the SPATIAL modifier variable within the literal es+*PRED-Cx and the precipitate Hagel 'hail'*

Word order / **SPATIAL modifier**	**PREC1** *hageln+es*	**PREC2** *es+hageln*	**Total**
LOCATION	2	7	**9**
GOAL	0	0	**0**
SOURCE	0	0	**0**
PATH	0	0	**0**
ZERO	0	16	**16**
COMBINATIONS	0	0	**0**
Total	**2**	**23**	**25**

Where a LOCATION is given, it is mostly an explicit place, such as cities, e.g., *in Wien* 'in Vienna', a larger area like *im Raum Schwabach* 'Schwabacher region', or a specific place within a city like *am Kölner Tanzbrunnen* 'at the Tanzbrunnen in Cologne', which is a live music venue with indoor and outdoor stages. One LOCATION in the wider context occurs with adverb *draußen* 'outside'. Another example of a setting provided in the wider context is given (219), where the LOCATION *bei der Kammgarn* does not directly modify the hail event.

(219) *Treffpunkt ist bei der Kammgarn. Sollte **es** "Krotten **hageln**", bleiben wir in der Kammgarn*.
(herthaglueck.at – 'herthaluck')
'Meeting point is at the Kammgarn. Should **it** '**hail** toads', we will stay in the Kammgarn.'

The six concordance lines that contain no SPATIAL modifier are all enumerations that list at least two different types of precipitation, as in (220).

(220) *Als **es** geregnet, gestürmt und **gehagelt hat**, habe ich im Geiste die kleinen California Poppies schon den Hügelbeethang runterschwimmen sehen* ... (metagarten.de 'metagarden')
'As **it** rained, stormed, and **hailed**, I pictured the little California poppies floating down the slope of the mound bed.'

In these enumerations, it becomes obvious that the speaker is referring to a whole spectrum of MEs, which emphasizes being hit by "all" the bad weather (notwithstanding the fact that hail seldom appears alone but is accompanied by rain or storm or thunderstorms) or that the weather was the extreme opposite, e.g., it changed quickly to sunshine, but that there is nothing "inbetween" (e.g., cast over but dry).

3.4.3.4 The NON-SPATIAL modifier variable (TIME, QUANTITY, QUALITY, STANCE)

The distribution and frequencies of the NON-SPATIAL modifier variable for literal meaning by word order are given in Table 92 .There are five concordance lines that do not involve any modifying

material (ZERO), including SPATIAL modifiers; another five utterances contain a single modifier only. Fourteen concordance lines contain a combination of at least two modifiers, NON-SPATIAL or SPATIAL.

***Table 92**. Distribution and frequencies of the NON-SPATIAL modifier variable for the* es*+PRED-Cx and the precipitate* Hagel *'hail'*

Word order/ MODIFIER	**PREC1** *hageln+es*	**PREC2** *es+hageln*	**Total**
TEMPORAL	2	12	**14**
QUANTITY	0	0	**0**
QUALITY	0	8	**8**
STANCE	2	8	**10**
Total	**5**	**33**	**38**
COMBINATIONS	2	12	14
ZERO	0	5	5
FIGURATIVE	98	77	175

The most frequent modifier is TEMPORAL (N=14), followed by STANCE (N=10), and QUALITY (N=8). There is not a single modifier involving QUANTITY. The individual modifiers are discussed in more detail in the subsequent sections.

Regarding the semantics of the TEMPORAL modifiers, there is no one category that crystallizes but a mix of deictics (*gestern* 'yesterday'), idiosyncratic expressions (*wir waren kaum zu Hause* als 'we were barely home when'), day of the week (*Sonntag* 'Sunday'), the time of day (*am Nachmittag* 'in the afternoon'), and so on. In contrast to other types of precipitation, hail—particularly in the *es*+PRED-Cx—does not appear with a modifier specifying the QUANTITY, but rather QUALITY. Accordingly, there is not a single concordance coded as QUANTITY, but there are eight concordance lines that contain some modifier of QUALITY or MANNER. In five utterances, the verb is used transitively and the QUALITY of the hail is specified as an object, i.e., *es hagelt Katzen* 'it hails cats' (N=2),[95] *Hunde und Katzen* 'dogs and cats', *Blitze* 'lightning', and *Krotten* ('toads', as in (219) above).[96] Twice, the MANNER in which it started to hail (and rain) is described as *plötzlich* 'suddenly'. Furthermore, QUALITY/MANNER is rather informally described as *quer* 'crossways' (which is also a path) and *in einer Tour* 'incessantly', which also contains an element of temporality.

[95] There are, in fact, 94 concordance lines returned for the simple concordance search *Katzen hageln* 'hail cats'. Only two of those 94 contain the collocation *Hunde und Katzen hageln* 'hail dogs and cats'.

[96] The word *Krotten* was unknown to me. I was not able to find it in a dictionary, neither was I able to find a direct translation of *Krotten*. It may be some dialect in Bavaria for *Kröten* 'toads', which seems the most reasonable option since there is an expression *es regnet Frösche* 'it rains frogs'. Take note that the example furthermore comes from an Austrian domain (.at).

A total of 10 concordance lines encompass expressions of stancetaking with six instances carrying negative semantic prosody, as in (221). There is one utterance with positive semantic prosody where the hail at the point of speaking/writing appears to be hypothetical ('it would hail cats', already mentioned above). In another two concordance lines the speakers claim being lucky because they just so evaded the hail or got home just when it started to hail, as in (222). In other utterances, STANCE is expressed in the left or right context, as in (223): *Glück haben wir auch noch.*

(221) *Dafür **war das Wetter wirklich sehr besch...eiden**, **es** regnete und **hagelte** in einer Tour!*[97] (carpeberlin.de)
'But **the weather was really lousy**, **it** rained and **hailed** incessantly!'

(222) *Wir waren kaum zu Hause als **es anfing zu hageln**, **Glück gehabt** **:-)*** (wouters-border-collie.de)
'We were hardly home when **it started to hail, lucky :-)**'

(223) *Und **Glück haben wir auch noch** ca. eine Stunde nach unserer Ankunft zieht ein Gewitter auf und **es fängt an zu hageln***. (timk.de)
'And **we're lucky**; about an hour after our arrival, a thunderstorm comes up and **it starts to hail**.'

(224) *(als) ich heute morgen um 8.30 Uhr gefüttert hab regnete es, als ich das Tor zum Grasstück hinterm Haus um 9.00 Uhr aufmachte **hagelte es (!)** und um 11.00 knallte die Sonne vom fast wolkenlosen Himmel….* (hufreheforum.de – hoofdeerforum)
'(When) I (did the feeding) this morning at 8.30 a.m. it was raining, when I opened the gate to the piece of grass behind the house at 9.00 a.m. **it was hailing (!)** and at 11.00 a.m. the sun was beating down from an almost cloudless sky.'

The expression of STANCE can be so minimal as to be reduced to the use of an exclamation point, as in (224), where the hail event is orthographically emphasized in such a manner as opposed to the other two weather events bracketing the hail event (rain in the left context and sun in the right context), which are not marked by additional punctuation. The next subsection is a brief excursion on semantic prosody within the overwhelming figurative usage found in the [*es+hageln*] returns.

[97] Literally, *bescheiden* translates to 'modest'. However, *bescheiden* is informally used as a substitute for the vulgar *beschissen* 'shitty', or *scheiße* 'shit', where people often drag out the homophone first part (*besch…*, as indicated by the ellipsis) before deciding which of the two words they will ultimately use, or if they even end up using *bescheiden*.

*3.4.4 Distribution and frequencies of variables across the three constructions (ARG+*fallen*, ARG+*geben, es*+PRED) with the precipitate* Hagel *'hail'*

In this section, the frequencies and distribution of variables by construction, i.e., the *es*+PRED-Cx and the two ARG-Cxx, as surveyed up to this point, are compared against one another. Mostly, the focus is on the two ARG-Cxx, since the [*es*+*hageln*] construction returned only 25 concordance lines that are being used literally. The two ARG-Cxx are very clearly and strikingly different in some regard, they certainly share characteristics as well. Having said that, even where they appear very similar, e.g., with respect to frequency counts, there are subtle difference regarding their semantics when the modifying material is examined more closely.

3.4.4.1 USAGE (LITERAL vs FIGURATIVE) and GENRE variables across the three constructions

The aggregate results of the different constructions with respect the distribution and frequencies of the USAGE variable are given in Table 93. There is a clear divide between the usage of the ARG-constructions and the *es*+PRED-Cx. While the ARG-Cxx are overwhelmingly used in the literal sense (N=190 and N=194), the *es*+PRED-Cx is predominantly used in the figurative sense (N=175), with only one-eighth of the data referencing actual hail events. Overall, this mixed result leads to more than two-thirds of the hail returns being used in a literal manner and roughly one-third being used figuratively.

Table 93. *Frequencies of literal and figurative USE by construction type (ARG+*fallen*, ARG+*geben*, and* es+*PRED) with the precipitate* Hagel *'hail'*

Construction /					**Relative Frequency**
USAGE	***Hagel+fallen***	***Hagel+geben***	***es+hageln***	**Total**	/409
Literal	190	194	25	**409**	**68%**
Figurative	10	6	175	**191**	**32%**
Total	200	200	200	**600**	**100%**

Table 94 below displays the distribution and frequencies of GENRE type by construction as well as the relative frequency of the GENRE across all three constructions. The results are furthermore displayed in Figure 28 (absolute frequencies) and Figure 29 (relative frequencies). The frequencies show what has been established individually: WEATHER TALK is the predominant genre type in the *deTenTen13* corpus, and, thus, across the constructions, although not as distinctly for

Hagel 'hail' as for the other precipitates. This becomes obvious particularly regarding the relatively mixed results within the [*Hagel+fallen*] data sets. Overall, less than half of the [*Hagel+fallen*] returns (N=189; 46%) are coded as genre type WEATHER TALK. Second most frequent are WEATHER REPORTS with less than one-third of the returns (N=119; 29%). The genre types METEOROLOGY and LITERARY are evenly distributed with 12% each, although the former returned one more concordance line than the latter. The least frequent genre type is OTHER with no more than two concordance lines. Figure 28 below visualizes the results in a bar graph, which nicely shows how, except for a single LITERARY [*es+hageln*] return, all three constructions are only present in genre type WEATHER TALK.

Table 94. *Raw and relative frequencies of the GENRE variable by construction type (ARG+fallen, ARG+geben, and* es*+PRED) and the precipitate* Hagel *'hail'*

Construction / GENRE	***Hagel+fallen***	***Hagel+geben***	***es+hageln***	**Total**	**Relative frequency** /409
WEATHER TALK	65	101	23	**189**	**46%**
WEATHER REPORT	46	73	0	**119**	**29%**
METEOROLOGY	37	13	0	**50**	**12%**
LITERARY	42	6	1	**49**	**12%**
OTHER	0	1	1	**2**	**5%**
Total	**190**	**194**	**25**	**409**	**100%**
FIGURATIVE	10	6	175	191	divided by 600 **32%**

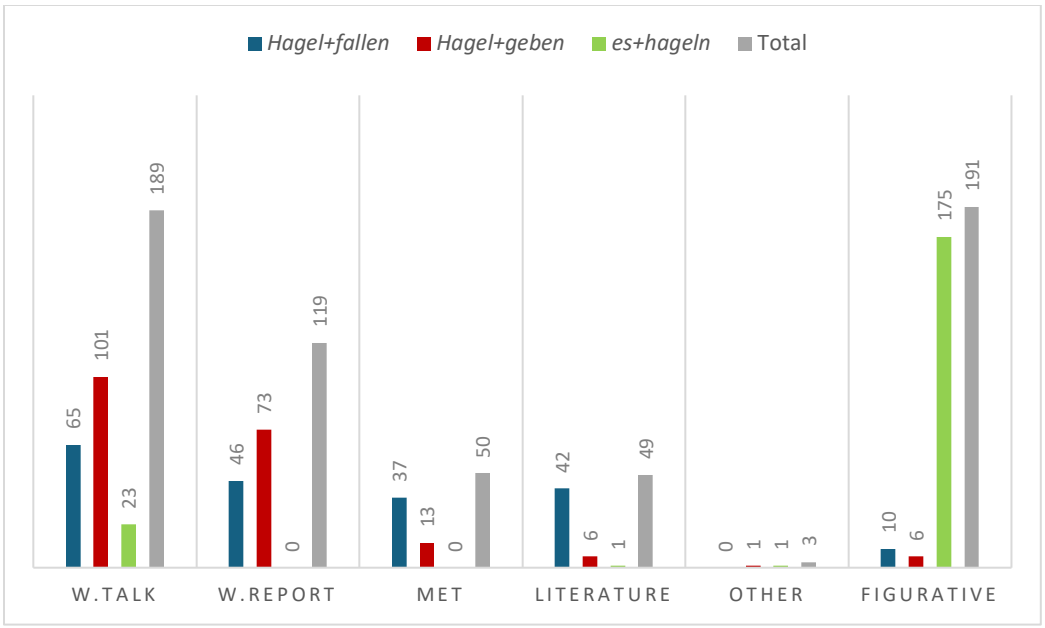

***Figure 28**. Distribution and absolute frequencies of the* GENRE *variable across all three constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED) and the precipitate* Hagel *'hail'*[98]

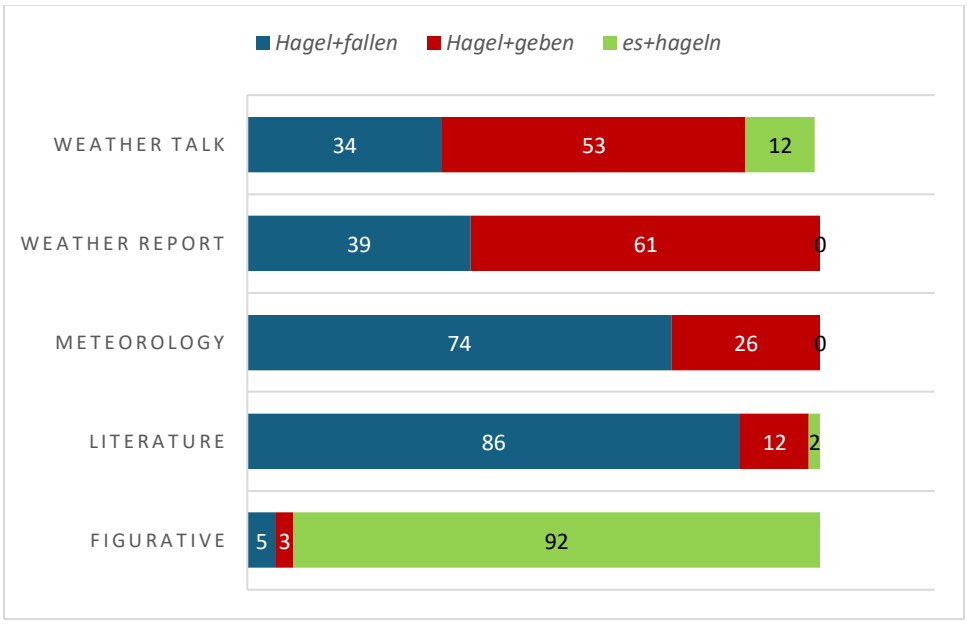

***Figure 29.** Percentual proportion of all three constructions (ARG+fallen, ARG+geben,* es*+PRED) and the* Hagel *'hail' precipitate to the individual type of* GENRE *(with "other" removed due to the low incidence)*

[98] WEATHER TALK and WEATHER REPORT are abbreviated as W. TALK and W. REPORT respectively, METEOROLOGY is abbreviated as MET, due to lack of space.

Figure 29 nicely visualizes how each type of genre clearly tends to attract different constructions. As already alluded to, the most balanced type of genre is WEATHER TALK, which occurs with all three constructions, but most frequently attracts the [*Hagel*+*geben*] construction. Obviously, since the [*es*+*hageln*] returns are so infrequent, the results are not conclusive. Since the majority of the [*es*+*hageln*] returns are coded as WEATHER TALK, it is likely that a "full data set" of 200 literal concordance lines would contain the highest incidence. However, it is clear, that the [*Hagel*+*fallen*] construction is not the chief contributor to this genre type but is more strongly attracted by genre types METEOROLOGY and LITERATURE. Interestingly, WEATHER REPORTS prefer the [*Hagel*+*geben*] construction over the [*Hagel*+*fallen*] construction although the [ARG+*fallen*] construction seems to be used in a more formal manner, generally speaking. It is difficult to make a meaningful statement about the [*es*+*hageln*] construction other than the stark predominance regarding figurative use.

3.4.4.2 The TAM (TENSE, ASPECT, MOOD) variable across the three constructions

Table 95 below shows the distribution and frequencies of construction by TAM level as well as the relative frequency across all three constructions. The results are furthermore visualized in Figure 30 (absolute frequencies) and Figure 31 (relative frequencies). Keeping in mind that the frequencies of the ARG-Cxx amount to more than 190 concordance lines each, while the *es*+PRED-Cx only yielded 25 literal results in total, each of three constructions overwhelmingly attracts PAST tense, a tendency that has crystallized for all precipitates across constructions. Almost two-thirds of the *Hagel* 'hail' data (62%) are coded as PAST tense, i.e., 254 concordance lines out of 409. The [*Hagel*+*geben*] construction makes up the highest share of PAST tokens with with more than half of the data followed by the [*Hagel*+*fallen*] construction with thirteen percentage points less. The remaining 5% are provided by the [*es*+*hageln*] construction. The second most frequent use of tense is EXISTENTIAL/GENERIC with one-fifth of the returns (N=77, or 19%). Here, the lion's share is taken by the [*Hagel*+*fallen*] construction (69%), which goes hand-in-hand with the majority of genre type METEOROLOGY attracting the [*Hagel*+*fallen*] construction. The remainder belongs to the [*Hagel*+*geben*] construction, since [*es*+*hageln*] only contains one concordance that is coded as GENERIC. Fittingly enough, the concordance in question linguistically describes the use of weather verbs in German. The [*Hagel*+*fallen*] construction also provides more than half of the concordance lines in PRESENT tense, which is the most infrequent TAM level overall. Of course, some of the EXISTENTIAL/GENERAL instances could arguably be analyzed as

habitual present tense. The remainder are fairly evenly distributed between [*Hagel+geben*] and the [*es+hageln*] construction, which is interesting considering the overall low incidence of the latter. The aggregate of "future intent" tokens is higher than that of PRESENT tense and the majority of these instances occur within the [*Hagel+geben*] returns. Slightly more than half of these occur in weather reports and encompass a modal, which explains their prevalence (all but one instance) in the PREC1 order. Within the "future intent" aggregate, the periphrastic FUTURE tense mostly occurs within the [*Hagel+fallen*] returns.

Table 95. *Raw and relative frequencies of the* TAM *variable across all three constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED) and the precipitate* Hagel *'hail' (literal instances only)*

Construction / TAM	*Hagel+fallen*	*Hagel+geben*	*es+hageln*	Total	Relative frequency /409
PAST	105	137	12	254	62%
PRESENT	20	8	7	35	9%
FUTURE_intent	12	26	5	43	10%
EXIS_GEN	53	23	1	77	19%
Total	190	194	25	409	**100%**

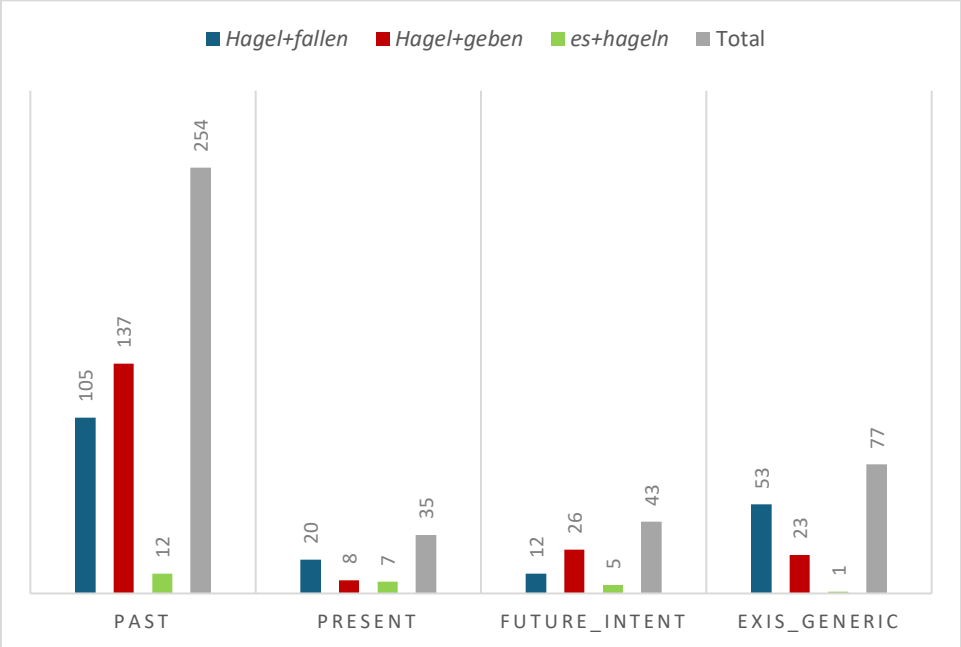

Figure 30. *Distribution and absolute frequencies of levels of the* TAM *variable across all three constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED) and the precipitate* Hagel *'hail'*

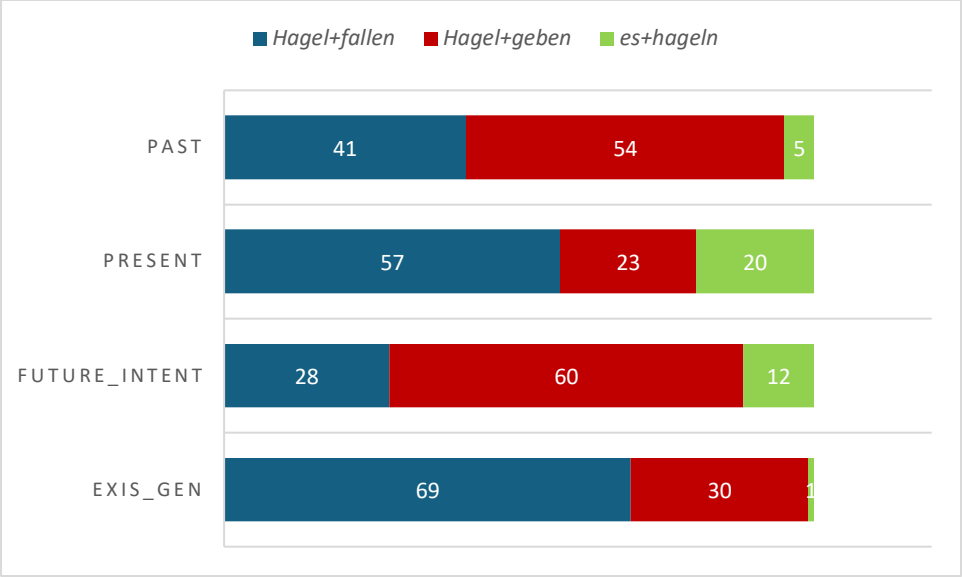

Figure 31. *Percentage of each individual construction (ARG+fallen, ARG+geben,* es*+PRED) per total of* TAM *level within the precipitate* Hagel *'hail' returns*

Due to the low incidence of returns for the [*es+hageln*] construction, it is difficult to make meaningful statements regarding the distribution, especially in comparison with the two ARG-Cxx. Having said that, it appears that the [*es+hageln*] construction tends to occur primarily with PRESENT tense and, to a less degree, with FUTURE_intent.

3.4.4.3 The SPATIAL *modifier variable across the three constructions*

The distributions and frequencies of the SPATIAL modifier variable for the ARG-Cxx and the *es*+PRED-Cx and the precipitate *Hagel* 'hail' are given in Table 96 and displayed in Figure 32. Keeping in mind that the *es*+PRED-Cx only contributes 25 concordance lines in total, the frequencies of the different SPATIAL modifiers are highly heterogeneously distributed between the three constructions.

Table 96. *Raw and relative frequencies of the* SPATIAL *modifier variable across all three constructions (ARG+*fallen*, ARG+*geben*,* es*+PRED) and the precipitate* Hagel *'hail'*

Construction / SPATIAL modifier	***Hagel+fallen***	***Hagel+geben***	***es+hageln***	***Total***	***Relative frequency*** /409
LOCATION	63	98	9	**170**	**41.5%**
GOAL	45	1	0	**46**	**11%**
SOURCE	46	0	0	**46**	**11%**
PATH	12	1	0	**13**	**3%**
ZERO	50	94	16	**160**	**39%**
Total	**216**	**194**	**25**	**435**	**106%**
COMBINATIONS	26	2	0	28	7%

The most frequent SPATIAL modifiers overall are LOCATIONS (N=170; 41.5%), followed closely by ZERO SPATIAL modification (160; 39%). GOALS and SOURCES are evenly distributed with (N=46; 11%). Lastly, PATHS amount to 3% of the data (N=13). Six percent of the data (N=25) overall contain a combination of at least two SPATIAL modifiers.

Surveying the data by construction paints a very different, more detailed picture as visualized in Figure 32. As the *es*+PRED-Cx is so infrequent, the focus is on the two ARG-constructions. The [*Hagel+geben*] construction can be described succinctly: SPATIAL modifiers are almost split between LOCATIONS and ZERO. The returns contain only two minor SPATIAL modifiers, one GOAL and one PATH. In contrast, SPATIAL modifiers within the [*Hagel+fallen*] returns are ubiquitous. Interestingly, here, too, LOCATIONS and ZERO modification are distributed fairly evenly. However, the "minor" SPATIAL modifiers GOAL and SOURCE are almost as frequent as zero modification; only PATHS have a comparatively low incidence.

The LOCATIVES differ semantically in that the [*Hagel+geben*] LOCATIVES consist mostly of "deictics" (N=32, such as *hier* 'here', *bei uns* 'here/at ours', *dort* 'there') and "locations" (*in Deutschland* 'in Germany', *zwischen Germering und Gilching* 'between Germering and Gilching') as well as some "cardinal directions" (*im Osten* 'in the east'). In comparison, the LOCATIVES in the [*Hagel+fallen*] returns encompass copious paths, especially sources and goals.

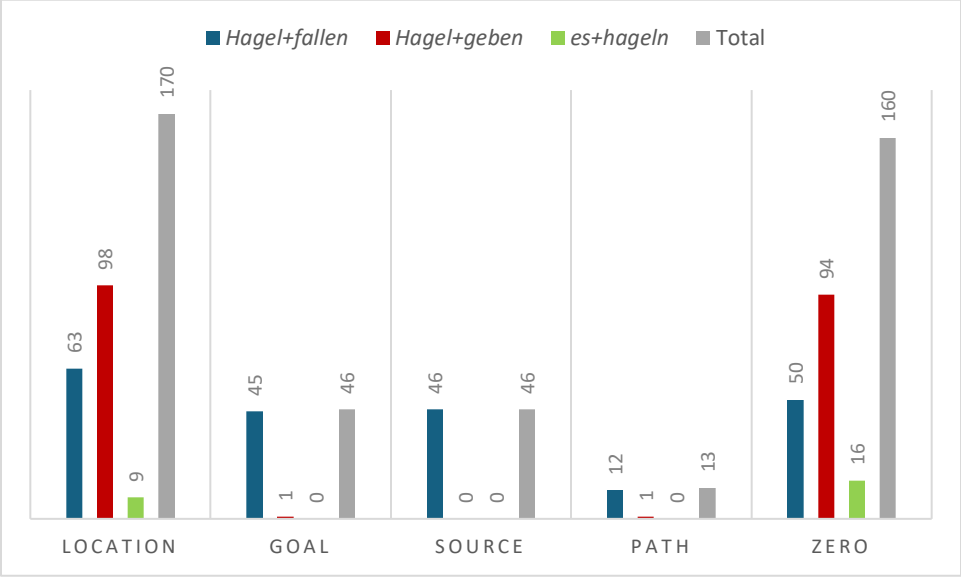

Figure 32. *Distribution and relative frequencies of the* SPATIAL *modifier variable across the three constructions (ARG+*fallen*, ARG+*geben*, es+PRED) and the* Hagel *'hail' precipitate*

Where there are "locations", they are also specific places, such as *in Bangkok*, *in ganz Ägypten* 'in all of Egypt' *or am Alpenrand* 'at the edge of the Alps', however, there are hardly any "deictics" (*vielerorts* 'in many places', *dort* 'there').

In sum, the [*Hagel+fallen*] construction is essentially almost solely responsible for all instances of SPATIAL modifiers that are not LOCATIONS: the GOALS, SOURCES, PATHS amount to one quarter of all literal corpus returns for the *Hagel* 'hail' precipitate. Regarding the [*es+hageln*] returns, most of the 25 utterances do not contain any SPATIAL modification whatsoever (ZERO), i.e., 4% of the data overall (N=16). There are nine LOCATIONS, of which four do not directly modify the ME. However, the *es*+PRED-Cx contributes too few concordance lines to generalize.

3.4.4.4 The NON-SPATIAL modifier variable across the three constructions

The distribution and frequencies of the NON-SPATIAL modifier variable across the three constructions are given in Table 97. The absolute frequencies are visualized in Figure 33 and the relative frequencies are given in Figure 34. SPATIAL modifiers are given in the last row, they contribute to the rows ZERO and COMBINATIONS. There are 30 concordance lines (7%), which do not contain any kind of modifier, SPATIAL or NON-SPATIAL. More than two-thirds of the concordance lines contain a combination of two different types of modifier, NON–SPATIAL and/or SPATIAL or. The most frequent type of NON–SPATIAL modifier is QUALITY, which occurs in almost half of the data (47%) and with one instance more than TEMPORAL modifiers (46.5%). Modifiers of STANCE are the third most frequent and occur in almost one-third of the *Hagel* 'hail' returns. The lowest incidence is provided by modifiers of QUANTITY, which only occur in 11% of the returns.

***Table 97**. Raw and relative frequencies of types of NON-SPATIAL modifiers across the three constructions (ARG+*fallen*, ARG+*geben, es*+PRED) and the precipitate* Hagel *'hail'*

Construction type	***Hagel+fallen***	***Hagel+geben***	***es+hageln***	**Total**	**Relative frequency**
NON-SPATIAL					/409
TEMPORAL	66	111	14	191	**46.5%**
QUANTITY	22	24	0	46	**11%**
QUALITY	107	77	8	192	**47%**
STANCE	55	64	10	129	**31.5%**
Total	**250**	**276**	**32**	**558**	**136%**
ZERO	6	10	5	21	5%
COMBINATIONS	131	136	14	281	69%
FIGURATIVE	10	6	175	191	divided by 425: 45%
SPATIAL	127	91	9	227	55.5%

Regarding the totals in Table 97 is surprisingly homogeneously distributed between modifiers of QUALITY and TIME which occur in more than half of the data. STANCE occurs frequently, too, i.e., in about one-third of the utterances. Only the QUANTITY modifier is trailing far behind. Having said that, the distribution by construction reveals a more diverse picture (cf. Figure 33).

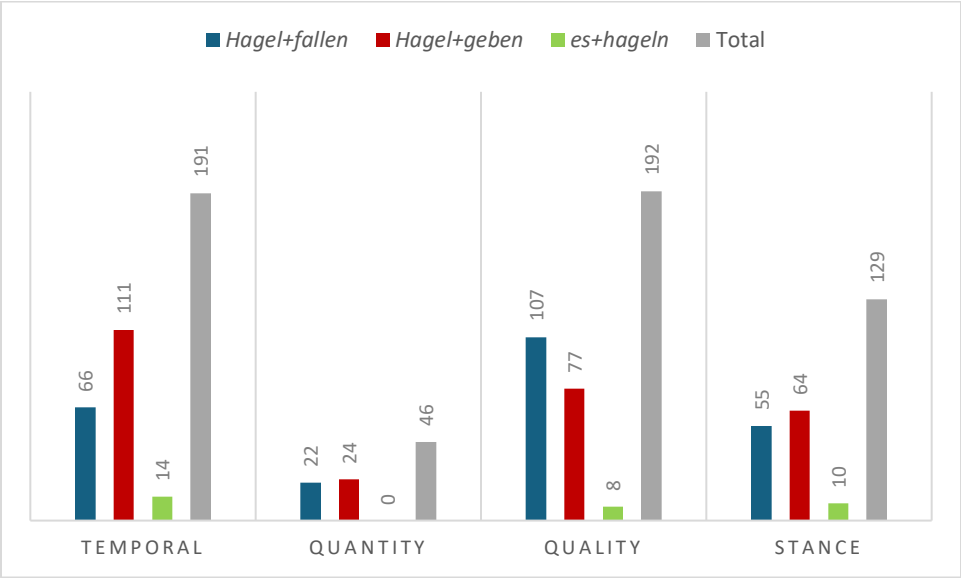

***Figure 33**. Absolute frequencies of the NON-SPATIAL modifier types TEMPORAL, QUANTITY, QUALITY, and STANCE across all three constructions constructions (ARG+fallen, ARG+geben, es+PRED) and the precipitate Hagel 'hail'*

Figure 34 shows how much each construction contributes proportionally, in percentages, to the total amount of type of NON-SPATIAL modifier. Again, the contribution of the [*es+hageln*] construction is of little consequence and Figure 34 mostly reveals differences between the two ARG-constructions. Having said that, even the low incidence of modifiers returned for the [*es+hageln*] construction demonstrates the lack of QUANTITATIVE modifiers and a focus on STANCE and TEMPORALITY with a slightly lower incidence for QUALITATIVE modifiers. At first glance, the differences between the two ARG-Cxx appear minimal. Especially, the distribution between modifiers of QUANTITY and STANCE is fairly even. The sharpest distinction is found for modifiers of TIME, which are much more likely to occur with the [*Hagel+geben*] construction with 25% more than [*Hagel+fallen*]. The second largest difference takes place with respect to the QUALITATIVE modifiers, which are more likely to occur with [*Hagel+fallen*] than with [*Hagel+geben*]. The semantic differences between the modifiers of each construction are elaborated on below.

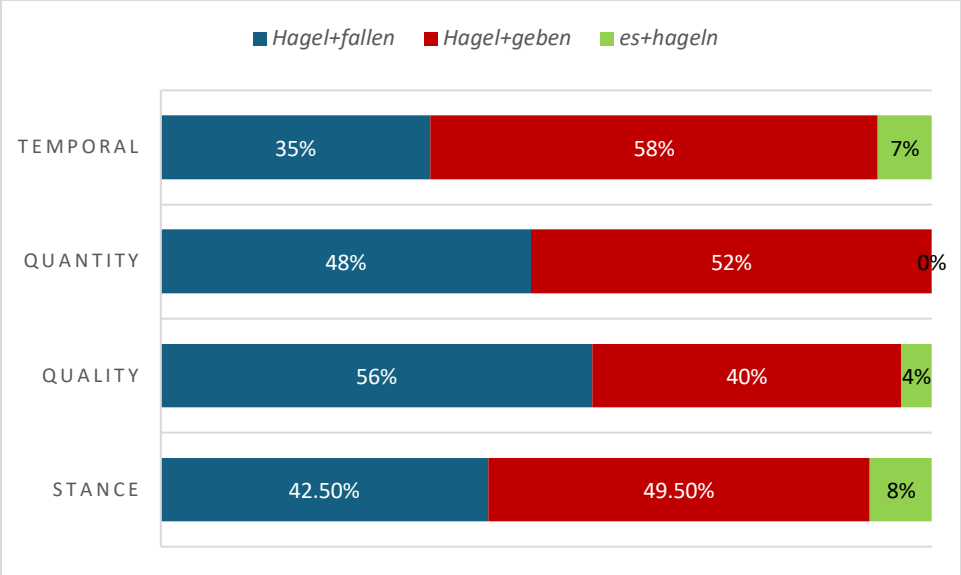

Figure 34. *Relative frequencies of the* NON-SPATIAL *modifier types* TEMPORAL, QUANTITY, QUALITY, *and* STANCE *across all three constructions constructions (ARG+fallen, ARG+geben,* es*+PRED) and the precipitate* Hagel *'hail'*

On closer inspection, there are obviously also differences with respect to the semantic and syntactic nature of the modifiers, which are discussed in the following paragraphs. The TEMPORAL modifiers in the [*Hagel+fallen*] returns tend to express "sequentiality" (*erst … dann* 'first … then', *später* 'later'), give "dates", either just the day (*am 14./15.* 'on the 14./15') or exact dates (*25.3.1907*; *1857*, *am 28. Juli abends 6 Uhr* '1857, on 28th July in the evening at 6 o'clock'), and introduce "time spans", especially with preposition *innerhalb* 'within' (*innerhalb weniger Minuten/kurzer Zeit/einer Stunde* 'within a few minutes/a short time/an hour'). The [*Hagel+geben*] construction indicates "time spans" as well but tends to use preposition *seit* 'since/for' (*seit dem Nachmittag* 'since the afternoon', *seit mehr als einem Jahrzent* 'for more than a century'). Furthermore, it tends to note the "time of day" (e.g., *mittags* 'at midday' or *in der Nacht* 'in the night') but not so much exact or full "dates". The [*Hagel+geben*] construction predominantly attracts adverbial "deictics" (such as *gestern* 'yesterday', *heute* 'today', *damals* 'back then', *diesmal* 'this time', N=25) and a variety of modifiers of "frequency" (*oft* 'often', *ab und zu* 'now and then', *erneut* 'again', *häufig* 'frequently', etc., N=18). Meanwhile, modifiers of frequency used in the [*Hagel+fallen*] returns appear to encode either high frequency (*oft* 'often' *öfters* 'more often, *immer wieder* 'again and again') or low frequency/absence (*selten* 'seldom', *niemals* 'never'). A type of temporal modification that occurs each once in the [*Hagel+fallen*] returns is by "months" and or "seasons"; The [*Hagel+geben*] construction employs both "seasons" (*im Sommer* 'in summer', *im Herbst* 'in autumn', N=6) but also a couple of "months"

(*im Mai/September* 'in May/September', N=2) more frequently. Additionally, both constructions contain several combinations and idiosyncratic modifications; [*Hagel*+*geben*] (N=21) more so than [*Hagel*+*fallen*] (N=10), but it contains at least 45 more modifiers of TIME than [*Hagel*+*fallen*] to begin with.

Modifiers of QUALITY for both ARG-Cxx are mostly descriptions of "size of the hailstones"; however, while the [*Hagel*+*geben*] constructionchiefly relies on giving the "size in centimeters" (N=30) and secondarily on describing the "size in words" (e.g. *kleinkörnig* 'small grained', *groß* 'large', N=19), The [*Hagel*+*fallen*] construction predominantly relies on adjectival modification (N=26) and much less frequently refers to the size in centimeters (N=12). Additionally, [*Hagel*+*geben*] more frequently refers to small size (*kleinkörnig* 'small grained', N=11) than large size (N=7). The opposite is true for [*Hagel*+*fallen*] where only three tokens refer to small sized hail. [*Hagel*+*fallen*] also contains eminently varied modification with respect to quality and manner of the hail stones or hailfall (N=36, e.g. *schwarz* 'black', *gelb* 'yellow', *fein* 'fine', *weich* 'soft', *plötzlich* 'sudden', *strichweise* 'scattered', *unerwartet* 'unexpected', *wie Steine* 'like rocks', *brennend* 'burning', *furchtbar* 'terrible', *eisig* 'icy', as well as 13 tokens referring to density) while the quality of hail in the [*Hagel*+*geben*] data primarily designates "force" (*leicht* 'light', *schwer* 'heavy', *heftig* 'severe', N=9) as well as a few "other" types (e.g., *dick* 'thick', *scheußlich* 'nasty', N=5). For either construction there are only few combinations of hail with other substances. *Hagel*+*fallen* contains 11 tokens of hail mixed with something else (e.g., *Blut* 'blood', *Feuer* 'fire', *Regen* 'rain') while the ARG *Hagel+geben* returns contain 12 tokens where *Hagel* 'hail' is the co-ME of *heftige Gewitter* 'severe thunderstorms' or other precipitation events (e.g., *Platzregen* 'downpour'). Lastly, the QUALITATIVE modifiers of the [*es*+*hageln*] constructionare fairly mixed. Having said that, in half of the data the verb is used transitively with an added object (*Krotten*/*Katzen*/*Hunde und Katzen hageln* 'hail toads/cats/cats and dogs').

With respect to expressions of STANCETAKING, it is almost exclusively negative, for either construction. There is only one utterance where the speaker regrets not having seen any hail during the current year. Where there is positive stance, it is usually accompanied by negation - these concordance lines are mostly found within the [*Hagel*+*geben*] returns with only a couple of instances in the [*Hagel*+*fallen*] returns. The [*Hagel*+*geben*] returns moreover contain several tokens of mirative *sogar* 'even' (N=6), drawing attention to the unexpectedness of the ME and the surprise element. Additional mirative stancetaking is conveyed by expressions such as *und was für ein* 'and what a (hail)!'.

Modifiers of QUANTITY are infrequent overall (11% of the QUANTITATIVE modifiers and 6% of all modifiers) and within the [*Hagel+geben*] returns are mostly provided by negation (ZERO, N=16) and varies otherwise (e.g., *im Überfluss* 'in spades', *den meisten* 'the most', ein *bisschen* 'a little'; N=8). Within the [*Hagel+fallen*] returns, it is also mostly "large quantities" (N=14) that are lexicalized, while "zero quantities" (N=8) and "small quantities" (N=6) only play a minor role. Moreover, the quantifiers in [*Hagel+fallen*] are visually evocative. There are five concordance lines that employ a *so viel, dass…* 'so much that…' construction: *… so viel, dass sich Straßen teilweise in reißende Bäche verwandelten* 'so much that some streets turned into raging streams', *…so viel, dass die Straßen komplett zu sind, dann man sogar Schwierigkeiten hat vom Fleck zu kommen* 'so much that the streets are completely blocked and you even have difficulty getting anywhere', *…so viel, dass er als Lawine einen Hang nieder kam und Straßen meterhoch bedeckte* 'so much that it came down a slope as an avalanche and covered streets meters high'. Meanwhile, the [*es+hageln*] returns do not contain any quantifiers.

3.4.5 Summary of Hagel *'hail' constructions*

The preceding quantitative and qualitative analysis of MEs employing the 409 returned concordance lines for the three constructions [*Hagel+fallen*], [*Hagel+geben*], and [*es+hageln*] has provided a comprehensive account of the constructional preferences of each of the three constructions and their interaction with a number of extra-propositional variables. Since the [*es+hageln*] construction returned predominantly concordance lines that are used figuratively, any conclusions have to be taken with a grain of salt. There appears to be a great deal of overlap between the two ARG-Cxx in some regard, however, each construction was also found to attract different variables to varying degrees or to show differences in lower-level semantics where frequency was not a differentiating factor.

I carried out a Pearson's chi-square test Table 93 (USAGE variable), Table 94 (GENRE variable), Table 95 (TAM variable), Table 96 (SPATIAL modifier variable), and Table 97 (NON-SPATIAL modifier variable). The distribution of literal and figurative utterances in Table 93 turned out to be statistically highly significant (X-squared = 428.59, df = 2, p-value < 2.2e-16). Unsurprisingly, the PRED-Cx is underrepresented regarding the literary usage and overrepresented regarding the figurative usage. The opposite is true for the two ARG-Cxx. For Table 94, the GENRE variable,

the frequencies are too low for a reliable application of the chi-squared test for a significant distribution. Nevertheless, where the predicate construction is removed, the distribution between the two argument constructions is statistically highly significant (X-squared = 52.433, df = 3, p-value = 2.422e-11). The same is true for Table 95, the TAM variable (X-squared = 26.335, df = 3, p-value = 8.113e-06). For Table 96, the SPATIAL modifier variable, the frequencies are too low for a reliable application of the chi-squared test for a significant distribution. Nevertheless, where the peripheral spatial modifiers (goals, sources, and paths) are combined, the distribution is highly significant (X-squared = 133.69, df = 4, p-value < 2.2e-16). For Table 97, too, the NON-SPATIAL modifier variable, the frequencies are too low for a reliable application of the chi-squared test for a significant distribution. However, the distribution turns out to be statistically significant where the predicate construction is removed (X-squared = 15.853, df = 3, p-value = 0.001215), where all three constructions are considered, but where the modifiers of quantity are removed (X-squared = 18.251, df = 4, p-value = 0.001102), and), and where all three constructions are considered, but where the modifiers of quantity and quality are combined (X-squared = 18.672, df = 4, p-value = 0.0009115). The more general observations are summarized in the following:

The [*Hagel+fallen*] construction is mostly attracted by genre type WEATHER TALK, but, to a lesser degree, is also prone to associate with the three other types of genre as well, specifically WEATHER REPORTS and the LITERARY genre. More than half of the data is inflected for PAST tense, mostly preterite, with little TAM marking, particularly in the PREC2 word order. EXISTENTIAL/GENERIC tense use is somewhat frequent as well, opposed to PRESENT tense use, which appears to be underrepresented. The [*Hagel+fallen*] construction attracts a variety of SPATIAL modifiers, especially GOALS and SOURCES, to a great extent and is almost solely responsible for the existence of these "minor" paths within the returns for the *Hagel* 'hail' data. As opposed to the [*Hagel+geben*] construction, there are hardly any deictics. Regarding NON-SPATIAL modification, the construction encompasses a considerable amount of qualitative modification where the size is mostly lexicalized by use of adjectives. The focus is on large size hail rather than small size. Lastly, there are several concordances which denote some other figure or "generic precipitation" falling *as hail*, such as water or water drops. Additionally, there are a few instances where *Hagel* 'hail' does not take center stage and appears as byproduct or co-ME to some other ME.

The [*Hagel*+*geben*] construction is mostly attracted by genre type WEATHER TALK as well but to a much higher degree than the [*Hagel*+*fallen*] construction. It is also frequently attracted by genre type WEATHER REPORT but only to small degree by the other two types of genre, especially LITERARY sources. The majority of the data, more than two-thirds, are inflected for PAST tense use. This is many more than the [*Hagel*+*fallen*] construction but just like in the [*Hagel*+*fallen*] returns, there is very little variation with regard to aspect marking, use of *Perfekt* or past perfect. Also, like [*Hagel*+*fallen*], there are very few instances of present tense. However, instances of "future intent" are somewhat frequent in the [*Hagel*+*geben*] returns. Regarding SPATIAL modification, the returned utterances are essentially split between contain LOCATIONS or ZERO modification but not much else. Regarding NON- SPATIAL modification, modifiers of TIME are the most frequent, especially deictics and expressions of frequency. Modifiers of QUALITY have the second highest incidence, but they are not as frequent as within the [*Hagel*+*fallen*] returns and also differ slightly in their make-up. They, too, focus on the size of the hailstones but do so by chiefly expressing the size in centimeters. Where the size is expressed in the shape of an adjective, the focus lies on the small size rather than largeness. Expressions of STANCE generally contain negative semantic prosody, which is true for all *Hagel* 'hail' returns (even the figurative ones are mostly negative); however, like all [ARG+*geben*] constructions surveyed up to this point, the [*Hagel*+*geben*] returns contain a lot of negation, which, in combination with stancetaking, acquires a positive meaning (*Gott sei Dank gab es keinen Hagel* 'Thank God it gave no hail'). Lastly, within the [*Hagel*+*geben*] returns, *Hagel* 'hail' often takes on the role of a co-ME rather than being the main event. As such, *Hagel* 'hail' is added as a PP, *mit Hagel* 'with hail', to the "main event". Most frequently, this main event is *Gewitter* 'thunderstorm' (*Gewitter mit Hagel* 'thunderstorm with hail').

Overall, the predominantly metaphorical use of the [*es*+*hageln*] construction left me with only about one-eights of the data of the other constructions to analyze. At the same time, the low incidence of the *es*+PRED-Cx and precipitate *Hagel* 'hail' prove how there can be no sweeping generalizations made about the predominant use of the predicate construction for precipitation events in the German language. In its literal sense, the [*es*+*hageln*] construction is predominantly used in more informal situations, i.e., WEATHER TALK, but not the more formal types of genres like METEOROLOGY, WEATHER REPORTS, or even in LITERARY sources. The harmful quality of *Hagel* 'hail' is also manifest in the type of modifiers, other than STANCE, occurring in the construction. While QUANTITY does not play much of a role (if any), the QUALITY of the

hailstones is of major importance, particularly with regard to their size. The TEMPORALITY, i.e., the point in time or the time span when the ME happened is stated fairly often but SPATIAL modification does not seem to be that prevalent. The most striking finding, apart from the predominance of figurative usage, is the variety of TAM marking within the small literal data set: each tense is present at least once, there is progressive marking in both PRESENT and PAST tense, use of *Perfekt* tense and past perfect, as well as instances of quasi-auxiliaries, which focus on the incipient phase of the ME.

3.5 Final summary of all three constructions with all four precipitates

The previous qualitative and quantitative analyses have highlighted the semantic and syntactic properties of each of the three constructions (ARG+*fallen*, ARG+*geben*, *es*+PRED) in combination with the four precipitates (*Regen* 'rain', *Schnee* 'snow', *Schneeregen* 'sleet', *Hagel* 'hail') and helped identify some general patterns, which are discussed in this subsection. Each construction was found to be associated with a specific set of variables, which are summarized in the following, keeping in mind that the *es*+PRED-Cx, due to the lack of [*es*+*schneeregnen*] and literal [*es*+*hageln*] instances, occurs with roughly 300 concordance lines less than the two ARG-Cxx. The first half of this section focuses on the differences between the construction types, while the second half of the section zooms in on the differences between the precipitate types. The relative frequencies that each construction type contributes to the respective variable (GENRE, TAM, SPATIAL modification, NON-SPATIAL modification) are given in Figure 35, Figure 36, Figure 37, and Figure 38 at the end of the first half of this summary and conclusion by construction type.

By Construction Type: Some general observations for the three constructions are that they predominantly associated with genre type WEATHER TALK. Obviously, the *deTenTen13* corpus was chosen as a proxy for spoken language, or at least one that was deemed to contain a higher incidence of informal WEATHER TALK than, for example, a pure newspaper corpus or similar. Therefore, the predominance of this genre type across constructions and precipitates is not surprising. Another observation, which is true for all three constructions, is that they mostly inflect for PAST tense. Again, this might have been expected since speakers, especially in a written medium, are more likely to report on events that have already happened than to

predominantly be concerned with potential weather events, wishes for certain weather conditions and/or events that take place at the moment of writing. Although the latter may emerge from certain sources, such as message board entries, comments, and or chats. In a corpus that contains weather reports only, we may chiefly find utterance inflected for what I call "future intent", i.e., the periphrastic future tense, present tense inflection that denotes future events, and, generally speaking, irrealis markers, such as modal verbs. Having said that, the results for the TAM variable are not quite as unambiguous as for the GENRE variable. Especially considering variation within the individual TAM levels, such as aspect marking, the data reveal subtle differences between the constructions. This is also true for the SPATIAL and NON-SPATIAL variable, all of which are further discussed in the following.

While the [ARG+*fallen*] construction is most likely to associate with genre type WEATHER TALK, it is also the least likely of the three constructions to be attracted by genre type WEATHER TALK. In turn, it is the construction most likely to be associated with the three other genre types, especially METEOROLOGY and LITERATURE, where it easily contributes more than half of the data, as shown in Figure 35, which gives the relative frequencies of each construction by genre type. This finding goes hand-in-hand with the tendency of the [ARG+*fallen*] construction to express EXISTENTIAL/GENERIC use with respect to the TAM variable, much more so than the other two constructions. In fact, [ARG+*fallen*] is twice as likely than [ARG+*geben*] to occur with EXISTENTIAL/GENERIC use. Having said that, it is still more frequently inflected for PAST tense and "future intent" intra-constructionally. Overall, [ARG+*fallen*] most infrequently occurs in PRESENT tense; furthermore, it does not contain many instances of quasi-auxiliaries, progressive marking, use of *Perfekt* or past perfect tense, which all appear to indicate immediacy, some sort of more direct entanglement with the ME, and a higher level of informality. None of these features are required for reporting on the weather in a certain region or, indeed, a weather forecast. Literary sources can, of course, account for any conceivable utterance. The most striking finding is apparent regarding the SPATIAL modifier variable: [ARG+*fallen*] is responsible for the vast majority of "peripheral" spatial modifications, namely, GOALS, SOURCES, and PATHS. It contributes roughly the same number of LOCATIONS as the [ARG+*geben*] construction and correspondingly the lowest amount of ZERO SPATIAL modification. This finding also dovetails the interaction of the GENRE and TAM variables since many of the METEOROLOGICAL and LITERARY utterances contain SOURCES and GOALS, specifying the origin of the precipitate and where it ends up, for example, with regard to the water cycle or how it affects the environment and/or an animate experiencer.

Lastly, the [ARG+*fallen*] construction, other than with modifiers of PLACE, frequently occurs with modifiers of TIME, which is true for all three constructions. Its greater formality is also mirrored in its increased unlikelihood to express stancetaking; it encompasses the fewest modifications of STANCE, but it does frequently lexicalize QUALITY notions as well as QUANTITY. The expression of the latter two appears to depend on the precipitate type, both with respect to frequency and semantics. For example, there are few QUANTITY modifiers within the *Schnee* 'snow' data sets but many within the *Regen* 'rain' returns and QUALITY modifiers with the *Schneeregen* 'sleet' precipitate focus on "lightness" while *Hagel* 'hail' MEs focus on (large) size. As already pointed out, the [ARG+*fallen*] construction appears to be used in more formal language, such as weather reports and forecasts, talking about the weather of a certain region (AVW), how weather works (HWW), and descriptions where the path or spatial notions are of particular importance, while the speaker's attitude towards the ME is secondary.

The [ARG+*geben*] construction most frequently associates with genre type WEATHER TALK, too, intra-constructionally but also across constructions. Furthermore, it is frequently attracted by WEATHER REPORTS. The construction associates with the genre types METEOROLOGICAL and LITERARY roughly only half as frequently as the [ARG+*fallen*] construction, although the proportions differ slightly by precipitate. With respect to the TAM variable, the [ARG+*geben*] construction is predominantly inflected for PAST tense (roughly two-third of the utterances intra-constructionally; N=408) as well as for "future intent" tokens (more than one-third of the utterances intra-constructionally, N=224) and contributes almost 50% for each TAM level across constructions, although the two ARG-Cxx are close together regarding the frequency with which they inflect for "future intent". This is probably due to the fact that the two ARG-Cxx provide the majority of WEATHER REPORT tokens. SPATIAL modification is straightforward within the [ARG+*geben*] construction and, generally, relatively evenly distributed between ZERO modification or a LOCATION. Having said that, as opposed to the other three precipitates, *Regen* 'rain' predominantly contains ZERO SPATIAL modification. Moreover, SPATIAL modifiers tend to occur in the shape of deictics. Except for modifiers of QUALITY, the [ARG+*geben*] construction contains the most tokens of all NON-SPATIAL modifiers. This was unexpected particularly with regard to expressions of STANCE, however, this might be due to the overall lower frequency of instances of the *es*+PRED-Cx. STANCE appears to be largely precipitate-dependent. It is mostly negative where the precipitates *Schneeregen* 'sleet' and *Hagel* 'hail' are concerned, but speakers bemoan the absence of snow rather than its presence. Absence is marked by negators

such as *kein* 'no' and the [ARG+*geben*] construction encompasses a striking amount of negation in comparison to the other constructions. Perchance, the generally received "existentiality" character of the [ARG+*geben*] finds its expression in negation with regard to MEs rather than in the expression of the actualized ME.

The data points from the [*es*+PRED] construction corpus returns are chiefly contributed by the two ARG-Cxx, since [*es*+*hageln*] is used predominantly in a figurative manner and [*es*+*schneeregen*] did not yield the amount of concordance lines aimed for due to the construction not being "standardized" yet. Despite the lower number of hits, the [*es*+PRED] construction contributes more tokens to the WEATHER TALK genre than the [ARG+*fallen*] constructions and roughly as many tokens to the LITERATURE and METEOROLOGY genres as the [ARG+*geben*] constructions (with a difference of 11 and 13 data points respectively). In contrast, it hardly associates with the WEATHER REPORT genre, which is interesting since *ipso facto*, weather reports talk about weather. However, Eriksen et al.'s (2010) presumed German default construction to talk about weather is the [*es*+PRED] construction. To be fair, weather reports are a special type of genre, yet their single intended purpose is discourse about the weather. In accordance with the paucity of genre type WEATHER REPORT, the utterances are seldom inflected for "future intent". Instead, they most frequently inflect for PRESENT, intra- and inter-constructionally. The returns for the [*es*+PRED] construction overall account for almost half of all PRESENT tense utterances across constructions and precipitates despite the "missing" concordance lines for precipitates *Schneeregen* 'sleet' and *Hagel* 'hail'. The returned concordance lines are slightly more frequently inflected for PAST tense than PRESENT tense (by nine tokens), but the incidences of PRESENT tense in the ARG-Cxx data sets are still much lower. Across all four precipitates, the [*es*+PRED] construction exhibits a large variety of TAM marking for progressive aspect, *Perfekt* tense, past perfect tense, and the use of quasi-auxiliaries, which focus mainly on the incipient phase of the ME. Across precipitates, the [*es*+PRED] construction appears more informal, dynamic, highly immediate and conveys the direct experience of the ME. In addition to the variety of TAM marking, this finding is also mirrored in the high amount of stancetaking. Despite the lower number of hits, the [*es*+PRED] construction, contributes more than one-third of all STANCE modifiers overall (cf. Figure 38), which is more than ten percent more than the [ARG+*fallen*] constructions. Here, too, modifiers of TIME are the most frequent, though. As opposed to modifiers found in the ARG-Cxx, in the [*es*+PRED] data sets they are often idiosyncratic and coincide with other events or activities the speaker is engaged in. Otherwise,

semantically, the temporal modifiers differ depending on the precipitate. For example, *Schnee* 'snow' MEs tend to express frequency and deictics while *Hagel* 'hail' MEs seem to prefer the expression of time spans. Modifiers of QUANTITY are overall rare within the [*es*+PRED] data sets. In fact, they are the most infrequent of all NON-SPATIAL modifiers across constructions. Modifiers of QUALITY appear to be infrequent, but the lack of modifiers for the *Schneeregen* 'sleet' and *Hagel* 'hail' precipitates makes it difficult to generalize since, within the sleet and hail precipitate returns, QUALITATIVE modifiers are generally more frequent for the [*es*+PRED] constructions than for the ARG-Cxx. With regard to SPATIAL modification, the [*es*+PRED] construction resembles the [ARG+*geben*] constructions in that it either mostly foregoes SPATIAL modification (ZERO) or tends to specify LOCATIONS. Having said that, as opposed to the majority of [ARG+*geben*] concordance lines, it also encompasses a few more peripheral SPATIAL modifiers such as SOURCES and GOALS.

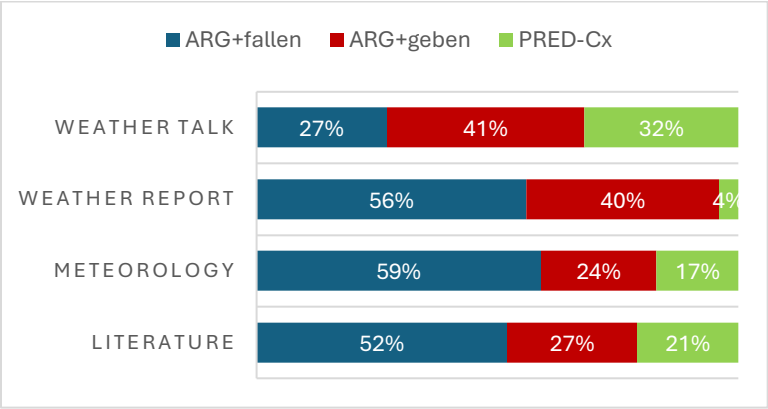

***Figure 35**. Relative frequency of construction type by types of GENRE in percentages*

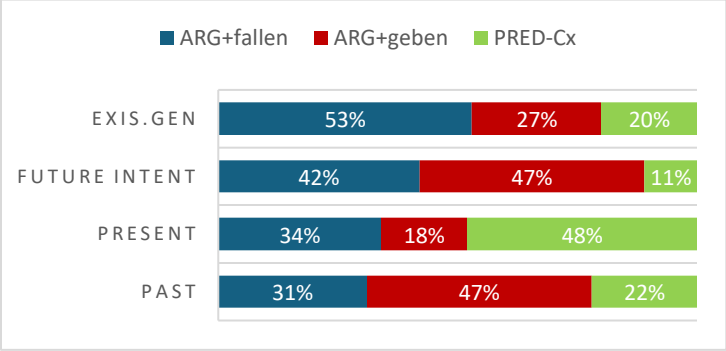

***Figure 36.** Relative frequency of construction type by TAM level in percentages*

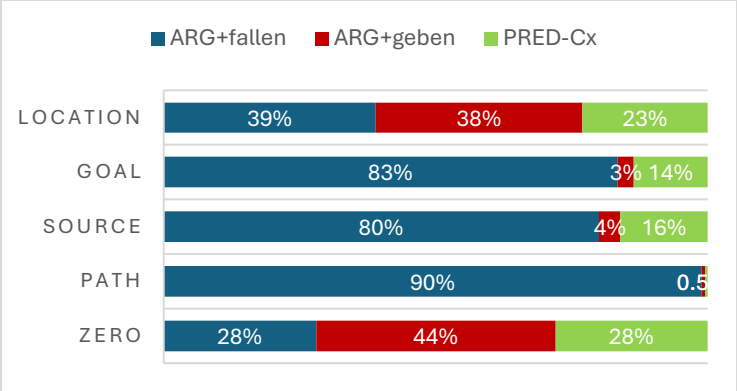

Figure 37. *Relative frequency of construction type by* SPATIAL *modifier in percentages*

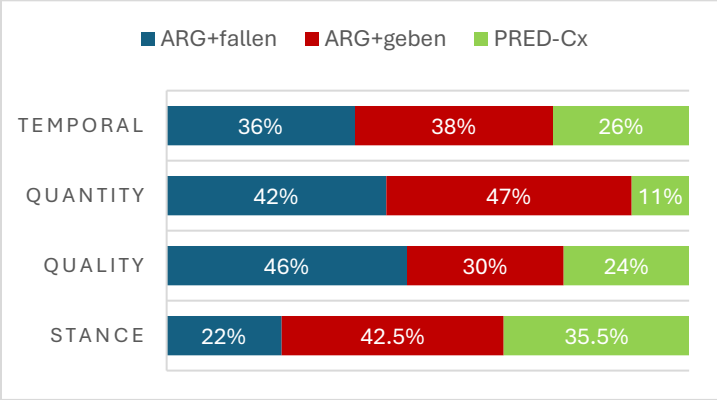

Figure 38. *Relative frequency of construction type by* NON-SPATIAL *modifier in percentages*

By Precipitate Type: If we look at the results by precipitate type, there appears to be a divide mostly between the more "standard" precipitates, *Regen* 'rain' and *Schnee* 'snow', and the less frequent, but possibly more salient, precipitates *Schneeregen* 'sleet' and *Hagel* 'hail'. The relative frequencies in percentages that each precipitate type contributes to the respective variable are given in Figure 39, Figure 40, Figure 41, and Figure 42 respectively, at the end of each paragraph.

Across precipitates, both sleet and hail associate the most frequent with genre type WEATHER REPORT. Rain and snow do probably not occur less in weather reports, but they associate with a higher incidence of other genre types, such as WEATHER TALK, which is why they appear underrepresented in Figure 39. Proportionally, they also contribute higher percentages to genre types METEOROLOGY and LITERATURE. This is true especially for *Regen* 'rain', while *Schnee* 'snow' is only slightly more frequent than *Hagel* 'hail'. *Schneeregen* 'sleet' barely contributes to the last-

mentioned genre types, especially METEOROLOGY. Since this genre type mostly deals with AVW situations, it explains why *Schneeregen* 'sleet' is so infrequent, as precipitation is usually noted as rain and snow is additionally listed in parts of a country/region where it frequently or occasionally occurs (if any), usually in combination with the time of year snow can be expected. As a "middle ground" precipitate between rain and snow, sleet is not a "default" precipitate, which warrants mention. That is to say, certain regions are known to be *rainy/pluvial* or *snowy*, but not "sleety". They are not "haily" either, but the METEOROLOGY instances for the *Hagel* 'hail' precipitate differ between constructions where [*Hagel+fallen*] mostly explicates how weather works and [*Hagel+geben*] refers to the weather of a region and the hail occurrence there, especially where hail is eminent.

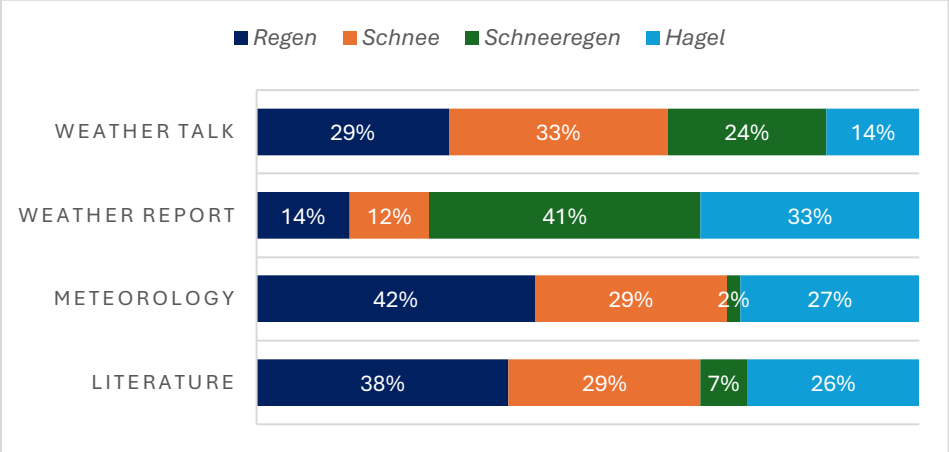

Figure 39. Relative frequency of precipitate type by type of GENRE in percentages

With respect to TAM marking, PAST tense turned out to be the most frequent TAM level across constructions and it is relatively evenly distributed between the precipitate types with roughly one-quarter each. Only *Hagel* 'hail' is more frequent and *Schneeregen* 'sleet' is less frequent than the other two precipitates. Again, it seems that reporting on a hail-ME that happened is more notable overall, possibly due to the rarity of the event. the negative impact it can have, or both. The reverse is true for *Schneeregen* 'sleet', which probably occurs less frequently naturally, to begin with, but is also not as impactful.

The overall higher proportion of "future intent" instances with respect to the *Schneeregen* 'sleet' precipitate probably goes hand-in-hand with the overall higher proportion of WEATHER REPORTS where, furthermore, *Schneeregen* 'sleet' almost always appears as a co-ME (mostly with *Regen* 'rain' or *Schnee* 'snow'). Interestingly, the *Regen* 'rain' precipitate

contributes the highest proportions of both GENRE type METEOROLOGY and TAM level EXISTENTIAL/GENERIC; these two levels of variables appear to co-occur generally. Again, where this is part of detailing the average weather situation in a certain country or region, it makes sense that it is common practice for rain and snow much more so than for hail and especially sleet. Having said that, the contribution of the *Hagel* 'hail' precipitate to EXISTENTIAL/GENERIC use is not that low (slightly less than one quarter of all tokens) and again, these are mostly instances concerned with how weather works.

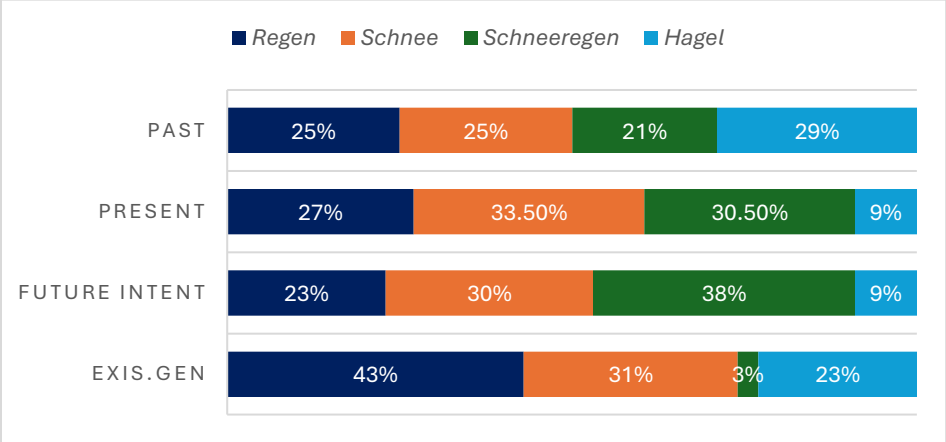

Figure 40. Relative frequency of precipitate type by TAM level in percentages

With respect to SPATIAL modification, we saw a clear preference of the [ARG+*fallen*] construction for the expression of peripheral SPATIAL modifiers, such as GOAL, SOURCE, and PATH while the [ARG+*geben*] construction was mostly split between LOCATIONS or ZERO modification. By precipitate, the results appear heterogeneously distributed as well. Here, it is the *Hagel* 'hail' precipitate, which proportionally contributes the majority of peripheral SPATIAL modifiers, followed by the *Schneeregen* 'sleet' precipitate regarding GOALS and SOURCES. Again, it is obvious that the goal of a hail-ME is of major importance, since hail tends to destroy things and, thus, where the precipitate ends up or what it ends up hitting, is important information. Furthermore, the majority of utterances containing a peripheral SPATIAL modifier come from HWW sources, which tend to ordinarily occur with these peripheral SPATIAL modifiers. LOCATIONS have proportionally the highest incidence within the *Schnee* 'snow' precipitate data sets. As pointed out earlier, for the corresponding concordance lines, it is often not clear (particularly without further context) whether the speaker refers to snow lying on the ground or falling snow, especially where the [ARG+*geben*] construction is concerned. However, for either option it

appears that the place the snow is "happening" is more important than its source, goal, or path, as it may turn into a genuine obstacle, especially where a figure (e.g., a person or a vehicle) has to find its way through the precipitate. A similar situation is not found for *Regen* 'rain' or *Schneeregen* 'sleet', which do not tend to leave traces or turn into obstacles unless their occurrence surpasses ordinary amounts. Following FRAME SEMANTICS, typical precipitation FRAME elements are time and place, however, the most typical precipitation type, *Regen* 'rain', surprisingly, contributes the majority of ZERO SPATIAL modifiers. It is possible that the SPATIAL modifiers appear in the wider context, that means a preceding or following sentence, which is not part of the KWIC returned by the corpus search. Still, the ostensible absence of SPATIAL modification is puzzling.

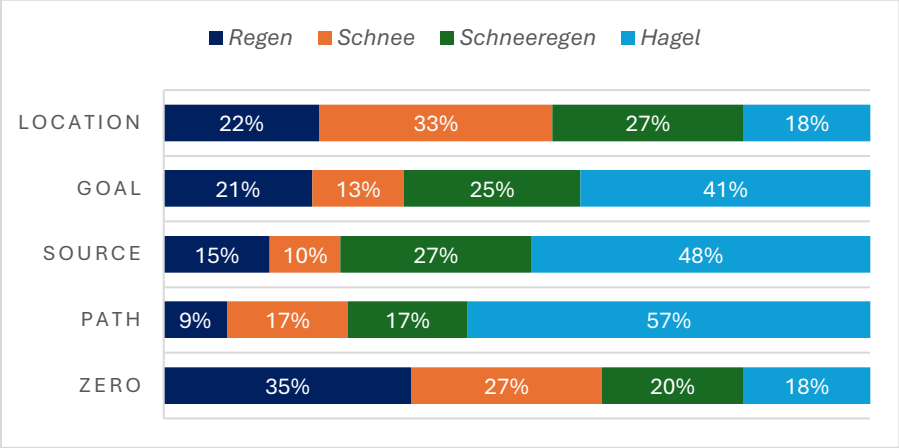

Figure 41. *Relative frequency of precipitation type by SPATIAL modifier in percentages*

Regarding the NON-SPATIAL modifiers TEMPORAL, QUANTITY, QUALITY, and STANCE, temporal modification is clearly the most frequent across precipitates (N=1248). It is, furthermore, a core element of the precipitation FRAME. Obviously, when or how long the ME takes place is of utmost importance. Accordingly, three precipitates contribute to temporal modification almost evenly with roughly less than one-third each, only *Hagel* 'hail' is half as frequent as each of the other three precipitates. Since hail events do not tend to be ongoing or long-lasting, as well as possibly not as easily forecastable, measuring them by other characteristics appears sensible. In fact, *Hagel* 'hail' provides the highest percentage of QUALITY modifiers overall despite the low number of concordance lines returned for the precipitate. Clearly, hail is more salient as a precipitate to begin with and often has destructive properties, which explains why there is a focus on its qualitative characteristics, which are predominantly lexicalized as size in

centimeters/numerals or size in adjectives such as *groß* 'big'. Similarly, this finding dovetails the higher overall proportional presence of peripheral SPATIAL modifiers, especially GOALS. As already mentioned, it is important "what got hit/destroyed/impacted" by the hail event, which in turn is contingent on the severity of the ME and the size of the hail stones (the bigger the more damage/impact). While both *Schnee* 'snow' and *Schneeregen* 'sleet' are less frequent overall and contribute similar percentages of QUALITY modification, *Regen* 'rain' still contributes more than one quarter of all QUALITY modifiers. Since the size of raindrops is not as crucial as the size of hailstones, qualitative characteristics of *Regen* 'rain' revolve around the lightness or severity of the event, where it depends on the construction type, which of the two it is; the ARG-Cxx tend to express the lightness while the *es*+PRED-Cx centers around the severity of the ME. Regarding the latter, there are enough synonyms for *regnen* 'rain' (V.), which encode the severe manner of the rainfall, such as *schütten*, *gießen*, *pladdern* (all can be roughly translate as 'pour') as well as synonyms for light rainfall, such as *nieseln*, *tröpfeln*, *fisseln* (all can roughly be translated as 'drizzle/drip'). Especially *nieseln* 'drizzle' is very common and it is possible that a completely different lemma (*es nieselte* 'it was drizzling') is more frequent than the modification of verb *regnen* 'rain' (*es regnete leicht* 'it was raining lightly'). Quality and quantity can appear very similar, which may explain why a precipitate with frequent expression of quality contains a low incidence of quantity and *vice versa*. Accordingly, the modification of QUANTITY appears to be more important than modification of QUALITY for the precipitates *Regen* 'rain' and *Schnee* 'snow', which contribute the highest proportion of any type of NON-SPATIAL modification across all precipitates. As we have already seen, there are slight semantic differences between the constructions, but apart from the [*Regen*+*fallen*] construction frequently employing numbers to express quantity (millimeters, liters), the ARG-Cxx more frequently lexicalize smaller quantities over large ones, or, in the case of the [*Regen*+*geben*] construction, contain negation. The majority of quantity modifiers occurs in the two ARG-Cxx while the *es*+PRED-Cx contains very infrequent QUANTITY modification. The same is true for the *Schnee* 'snow' precipitate; however, here, larger quantities are more frequently expressed than smaller quantities and the [*Schnee*+*geben*] construction encompasses a high amount of negation. By and large, with regard to the aggregate frequencies, modifiers of QUANTITY and QUALITY are surprisingly evenly distributed with 502 and 504 tokens respectively, despite the heterogenous distribution by precipitate.

Lastly, the contributions of each precipitate to STANCE are the most even of all NON-SPATIAL modifiers, surprisingly so. The highest proportion is provided by the *Schnee* 'precipitate'. The lower contributions of the precipitates *Schneeregen* 'sleet' and *Hagel* 'hail' can probably be explained by the overall lower number of concordance lines. Obviously, the proportional frequencies do not indicate whether STANCE is supplied in the shape of positive or negative semantic prosody. For all precipitates, negative semantic prosody appears to be prevalent. Where there is positive semantic prosody, it is often combined with negation, so that the non-actualized ME is perceived as positive. The only exception here is the precipitate *Schnee* 'snow', which is often anticipated, wished for, or just in general a joyful event. In turn, the lack of snow is often bemoaned. Expression of stancetaking is frequent. In fact, STANCE is the second most frequent NON-SPATIAL modifier (N=660), which seems intuitive for data sets that mostly associate with genre type WEATHER TALK. Due to the fact that the weather greatly influences our day-to-day lives and plans, it seems obvious that speakers not just talk about the weather but also voice their evaluations about the weather situation. Since this paper treats precipitation rather than "good weather", it also appears intuitive that the majority of stancetaking is negative rather than positive, with the exception of snow events.

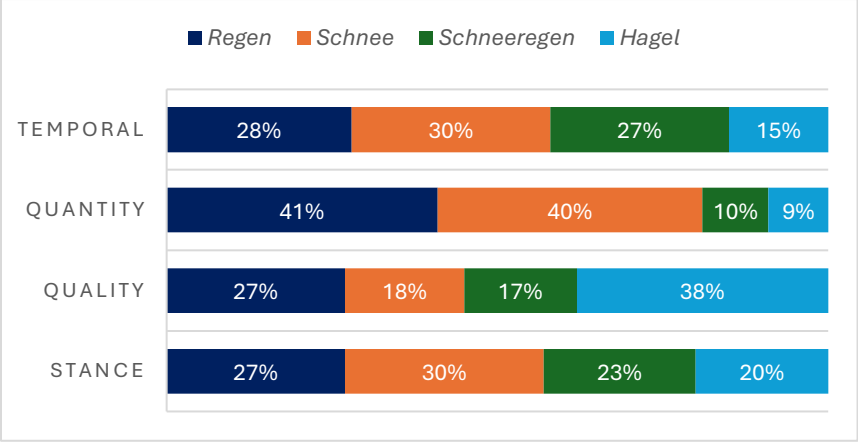

Figure 42. *Relative frequency of precipitate type by NON-SPATIAL modifier in percentages*

The preceding quantitative and qualitative analysis of precipitation events has provided a comprehensive account of mostly constructional preferences of each of three constructions across four precipitate types as well as some of the preferences by type of precipitate. Since the *es*+PRED-Cx in combination with the two precipitates *Schneeregen* 'sleet' and *Hagel* 'hail' only returned a small number of concordance lines that could be analyzed, the results for these precipitates must be taken with a grain of salt. Having said that, their paucity, of course, also

revealed certain preferences including where the respective construction or precipitate returned a high incidence despite the overall lower returns. While two constructions or precipitates often appear to have similar incidences, with the third and/or fourth one being the odd one out, the analysis of the variables and their different levels revealed that each construction often associates with a similar set of variables and, accordingly, tends to have a particular profile regarding its usage and features.

The next chapter surveys intransitive and transitive constructions, which are marked in that there are mostly two elements, which encode the weather event equally or each one facet of the ME. Namely, this is the ARG-PRED type as identified by Eriksen et al.'s (2010) typology where the semantic load is shared between the argument and the predicate (*Regen regnet* 'rain is raining'). Additionally, I examine prepositional constructions (*bei Regen* 'when/while it's raining', lit. 'at rain'), the use of causal agentives such as *Frau Holle* 'Mother Hulda' and *Petrus* 'St. Peter', where the former is usually associated with snowfall and the latter with rainfall or dry weather (when he "means well"). Furthermore, I explore other idiosyncrasies, such as *-wetter* '-weather' compounds, which showcase the astounding variety of weather talk a speaker has at their disposal.

CHAPTER 4

Marked Weather Constructions

4.1 Introduction

As several scholars have pointed out (Ruwet 1986: 365, Langacker 1991, Eriksen et al. 2010 and 2012, etc.) weather constructions are "ill-behaved", "weird", or simply highly marked to begin with, mostly due to the fact that they do not contain proper participants, which are independent of the event itself. Accordingly, speakers are left with the task to shoehorn a conceptually difficult event—where the participants are not always easily discernible (viz. the easily discernible shining sun as the star at the center of the Solar System vs. more diffuse clouds or the sky being overcast)—into a grammatically well-behaved structure. The intransitive expletive subtype of the predicate construction (*es regnet* 'it rains') is one such attempt, which, typologically speaking, is very rare outside of Europe and typical of Northern and Western European languages (Eriksen et al. 2010: 574–575). It represents an unmarked sentence structure imposed onto a marked event, which, uncharacteristically for the German language, conflates a figure with a motion event (cf. fn.4)

Chapter Three concerned these paradoxically "unmarked" constructions in German, i.e., the intransitive predicate construction (*es regnet* 'it rains') as well as the intransitive argument construction (*Regen fällt* 'rain falls'), which both follow the formula of one element being semantically heavy and encoding the ME while the second element is semantically light and more or less redundant or only syntactically required. In the intransitive predicate construction, this light element is the "expletive" pronoun *es* 'it' and in the intransitive argument construction this is the light verb *fallen* 'fall'. The third construction, [*es*+*geben*+ARG_{ME}]—and to a lesser degree [*es*+*fallen*+ARG_{ME}]—is even more odd in that it has two semantically light elements, but it is unmarked in that it is a productive transitive construction, which is generally used to indicate the existence of something. Eriksen et al. (2010) consider the expletive predicate type to be the default construction used to talk about precipitation in German. This may well be true if we consider the raw frequencies of each construction type only;[99] however, the analysis of the

[99] The raw frequencies for the *Regen* 'rain' precipitate were 228,177 instances for the PRED-Cx, 16,860 instances for the ARG+*fallen*-Cx and 2,915 instances for the ARG+*geben*-Cx (see Chapter 3).

extra-propositional content, which Eriksen et al. (2010) expressly disregard in their study (modification not considered) tells a different story. Weather reports, for example, the genre dedicated to addressing nothing but the weather, do not frequently employ the predicate construction. Incontestably, context matters, and Chapter Three revealed that each construction associates with a specific set of variables depending on their usage. Chapter Four now treats similarly syntactically unmarked intransitive and transitive constructions, however, the intransitives are marked in that there is no single semantically light element, but both the argument and the predicate are responsible for encoding some facet of the ME.

This chapter first discusses intransitive clause types (Section 4.2.1) that belong to Eriksen et al.'s (2010) argument-predicate type, namely, the "cognate type" and the "split type". The cognate type (Section 4.2.1.1), is basically a special case of either the intransitive argument type or the intransitive predicate type, with the difference being that no single element is responsible for the semantic encoding of the ME, but both elements express the same facet of the ME, e.g., *rain is raining* or *snow is snowing*, seemingly tautologically so. It is not clear to me whether there is one language that predominantly uses this type or whether some languages mostly use the cognate type for the rain event. Where the verb means 'to rain' it still does not specify the precipitate type, which has to be mentioned separately (except for the "default" precipitate type rain, which is implied). In the split type (cf. Section 4.2.1.2), both elements contribute to the ME by encoding different aspects of the same event. These can further be divided into which element is chiefly responsible for encoding the ME, so that there are again the two subtypes: the argument type and the predicate type. Neither the cognate type nor the split type are the default choice for everyday MEs in the German language, but they are marked, valid options with respect to specific situations and preferentially associate with certain genres. The latter is particularly true for the cognate type.

Section 4.2.2 pertains to transitive constructions that express MEs. In Chapter Three, and particularly within the returns for the [ARG+*fallen*] construction, the subject of intermediaries or agents that *let snow fall* already surfaced. For the most part, these are non-volitional entities like the sky, clouds, or the atmosphere. These more diffuse meteorological elements are obviously integral and necessary parts of the event itself, but they do not have to be obligatorily mentioned as the source of the precipitate (nor are they to a large extent), which is implied or part of a speaker's world knowledge. In this section, the focus lies less on these natural participants but more specifically on the recruitment of religious or mythological figures

that act as the causal, volitional agent behind the ME. While this agent can generally be associated with some unspecified *Wettergott* 'weather god', the two figures that loom large here are *Petrus* 'Saint Peter' and *Frau Holle* 'Mother Hulda'. These are, as opposed to atmospheric conditions, obviously not intrinsic parts of the ME, but extrinsic participants (cf. Eriksen et al. 2010: 570). They are, on the one hand, so ubiquitous that their invocation does not seem particularly marked. On the other hand, they are so culturally specific that the knowledge of their connection with MEs must, at some point, be explicitly acquired. For *Frau Holle*, this is primarily true with regard to the *Schnee* 'snow' precipitate and for *Petrus* mostly, but not exclusively, with regard to the *Regen* 'rain' precipitate. Generally speaking, the marked constructions treated here appear to not be equally available for every precipitate type, which makes them even more marked.

Sections 4.3 and 4.4 treat other construction types not necessarily part of Eriksen et al.'s (2010) typology. In 4.3, I address the "prepositional type" where the ME, happening simultaneously as some other event while not being particularly in focus, is introduced by a preposition (*bei* 'at'). These MEs may refer to any type of weather (*bei Wind* 'at wind', *bei Regen* 'at rain', *bei Sonnenschein* 'at sunshine') specifically, but also more generally (*bei gutem/schlechtem Wetter* 'at good/bad weather'). In 4.4, I turn to very idiomatic expressions of weather, such as *Kaiserwetter* 'Emperor weather' or *Fritz-Walter-Wetter.* The discussion of these idiomatic expressions is followed by with a brief overview of some specific regional expressions (4.4.3) as well as intriguing examples or combinations of ME constructions and precipitate types that I came across while annotating the data frames (4.4.4). Last but not least, the review of figurative examples in Section 4.5 brings us full circle to the data frames presented in Chapter Three, which revealed distinct tendencies for construction and precipitate types to occur metaphorically, rather than literally.

4.2 Marked Argument-Predicate Constructions

In German, marked ME constructions having intransitive structure involve two semantically heavy elements. Both the argument and the predicate express different facets of the ME. Section 4.2.1 considers the use of the cognate type in the German language, followed by the use of the split type in Section 4.2.2.

4.2.1 The cognate type

The cognate type is neither the default ME type nor very frequent in German. It is, however, not difficult to find examples of cognate ME constructions, be they highly marked utterances either in literary genres, like poetry, but also in everyday language. The different manifestations of the cognate type are detailed in this section and supported with concordance lines from the *deTenTen13*, the default for examples, and the *DWDS* corpora, and examples from *Google* (where only the latter two sources will be indicated where applicable). The cognate type is, furthermore, frequent enough to be captured in the overview discussion about the verb *regnen* 'rain' in the *DWDS* dictionary. The *DWDS* entry gives three usage types subsumed under item #1 with at least one example per usage type. First of all, it mentions the more canonical usages of the verb *regnen* 'rain', where the first, *es regnet* 'it rains', is paraphrased with *Regen fällt* 'rain falls'. The second usage type is given as the transitive *es regnet etwas* 'it rains something', exemplified by *es regnet große Tropfen* (=*in großen Tropfen*) 'it rains big drops (=in big drops)'. The third usage type is given as the cognate type construction *der Regen regnet* 'the rain rains', which is preceded by the frequency designation *selten* 'rare'. The following example is a line from Otto Nicolai's opera *Die lustigen Weiber von Windsor* 'The Merry Wives of Windsor': *der Regen regnet jeglichen Tag* 'the rain it raineth every day'.[100] I elaborate on this utterance, which is based on a play by Shakespeare, later in this section, as it was returned from a *DWDS* search query with several tokens.

In contrast to the *regnen* definition, the *DWDS* dictionary does not return cognate type entries for the other precipitates, namely *Schnee* 'snow', *Hagel* 'hail', or *Schneeregen* 'sleet'. Moreover, it is quite apparent, from various corpus search queries as well as dictionary definitions, that there are stark differences between the precipitate types and their propensity to occur in a cognate construction. Table 98 shows the different frequencies obtained for the respective cognate constructions by precipitate as queried in the *deTenTen13* corpus and the *DWDS* corpora. Clearly, the precipitate *Regen* 'rain' returns the highest incidence of meteorological cognation in both corpora. The *DWDS* corpora search query returned more than twice as many tokens as the *deTenTen13* query, which is likely due to a token from a German 19th century ballad that appears to have been reproduced in several textbooks and is

[100] Source: www.dwds.de/wb/regnen (last accessed February 5, 2024)

responsible for roughly one-fourth of all of the returns. A higher incidence of *Schnee schneit* 'snow snows' returns is found in the *DWDS* corpora (N=53) than in the *deTenTen13* (N=23) as well. Instances of *Hagel hagelt* 'hail hails' (N=2) are only returned from the *deTenTen13.* There are no instances of *Schneeregen schneeregnet* 'sleet sleets' in either corpus. It should also be noted that the majority of utterances appear in V2 word order; less than 10% of all returns give the cognate verb before the noun. The search query in the *deTenTen13* was carried out by means of a basic search with no gap between item #1 and item #2 (e.g., [word="Regen"][lemma="regnen"]) since a CQL query similar to those in Chapter 3, which allowed for up to eight items between the noun and the verb, mostly rendered false returns. The search query for items in the *DWDS* corpora was simply *Regen regnet*, *Schnee schneit*, etc., since a query containing two ampersands (e.g., *Regen* && *regnet*) returns any sentence with both of these items, in some shape or form, without them constituting a cognate construction. I then went through each of the *DWDS* corpora individually and added together the number of returns for the total count in Table 98 and removed any double entries that came from the same source.

Table 98. *Frequencies of cognate constructions by precipitate in the* deTenTen13 *corpus and the* DWDS *corpora*

Precipitate		***deTenTen13***	***DWDS***	**Total**
Regen regnet	'rain rains'	**75**	**184**	**259**
Schnee schneit	'snow snows'	**23**	**53**	**63**
Hagel hagelt	'hail hails'	**2**	**0**	**2**
Schneeregen schneeregnet	'sleet sleets'	**0**	**0**	**0**
Total		**100**	**237**	**337**

The returns can be grouped into different categories depending on their semantic make-up, which, in turn, also differ according to the precipitate type. The remaining discussion in this section deals with *Regen* 'rain' and *Schnee* 'snow' only, since the other precipitates yielded hardly any results. Rather than treating the precipitate types separately, I will go by the different manifestations of the constructions and provide appropriate examples. We begin with the utterances which are the most unmarked or commonplace in that they encode actualized MEs happening in trivial WEATHER TALK situations. These are neither tautological utterances, necessarily true and apparently uninformative, nor are they instances of highly marked poetic language. That being said, each of the examples differs in its construal. As already remarked on in Chapter 3, the weather conditions at the time of an outside sports event, like a match or a race, are critical as they do not only impact the athletes' performance, but may potentially

endanger the athlete. Accordingly, reports on sporting events generally contain some information on the concurrent weather situation. However, there is obviously no need to resort to a highly marked construction, like the cognate construction, where the predicate construction would suffice. However, as we saw in Chapter 3, the predicate construction often occurs with quasi-auxiliaries, which focus on the incipient or terminative phase of the ME in relation to a particular phase of the sporting event, so in a way, the cognate construction is more direct and immediate as it requires less overt TAM marking.

The example in (225) reports on a bicycle race, where the speaker easily could have employed the predicate construction, *es regnet jetzt* 'it is raining now', instead of a cognate construction.

(225) ***Der Regen regnet jetzt****, aber George Bennett sieht immer noch stark aus, als er auf die Linie zu klettert.* (Source: Google, racecentre.velon.cc)
'**The rain is raining now**, but George Bennett still looks strong as he is climbing towards the (finish) line.'

The temporal frame is specified by the adverb *jetzt* 'now', which implies progressivity, but also that it did not rain earlier in the race. The subordinate clause is introduced by an adversative conjunction, *aber* 'but', which juxtaposes the difficult weather conditions with the athlete's physical strength. Example (226), too, is from a report on a sports event; however, here, it is a soccer match that the speaker follows.

(226) ***Der Regen regnete****, die Nationalmannschaft schleppte ihr 1:0 mühsam durch die zweite Halbzeit.* (frawemedia.de 'Frank Weber Media')
'**The rain was raining/rained**,the national team schlepped their 1:0 laboriously through the second halftime.'

As opposed to (225), the construal in (226), brought about by the stringing together of two main clauses, the lack of a temporal modifier like *jetzt* 'now', as well as the choice of manner words like *schleppen* 'drag/schlep' and *mühsam* 'laboriously' impart a sense of a lengthy, effortful event where the action on the playing field is as unpleasant as the accompanying weather. In this manner, the choice of the cognate construction appears more deliberate and intentional in the conceptualization of the ongoing event. The same is true for (227), where the construal parallels that in (226) and evokes a similar mood or atmosphere.

(227) ***Der Schnee schneite****, die Füße froren, und der Weihnachtsmann hatte schon erste Eissterne in den Barthaaren, als wir endlich den Parkplatz betraten (…)* (mimimueller.de)
'**The snow was snowing**, the feet were freezing, and Santa Claus already had the first ice crystals in his whiskers when we finally entered the parking lot.'

Although these two utterances were more obviously prepared as an edited-to-be-read text (they are both blog entries rather than more conceptually oral internet forum messages), they still represent instances of WEATHER TALK rather than instances of the LITERARY genre. Take note that all three examples in (225)–(227) contain the definite article *der* 'the.MASC', which is often an indicator that the precipitate may have been introduced earlier in the discourse, often in the form of a more standard predicate construction. That said, this earlier introduction of the precipitate is not the case for examples (226) and (227). The link for (225) is broken so that the context cannot be verified. For (228), the target ME is preceded by *Es war ein stark weißer Schneetag* 'It was an intense white snow day' so that the 'snow day', preceded by the indefinite article *ein* 'a', is introduced before the subsequent sentence follows with a definite article before the precipitate in a cognate construction.

(228) *Es war ein stark weißer Schneetag.* ***Der Schnee schneite*** *pausenlos.*
(gsv-zwickau.de - 'Deaf Sports Club Zwickau')
'It was an intense white snow day. **The snow snowed** pauselessly.'

The target ME itself is modified for the MANNER in which the snow is snowing: *pausenlos* 'pauselessly'. This utterance, too, is from a report on a sports event where the participants of a bowling tournament decided to take the train rather than a car due to the whiteout conditions. In (229), taken from an online review that a customer left about a tire model, the speaker introduces the precipitate by pointing out that he did not expect it, but was impressed with the car tires for which the weather conditions posed no problem.

(229) *Zuerst bin ich mit diesen Reifen ins Gebirge gefahren, wo unerwartet* ***Schnee schneite****. Die Reifen sind mit diesen Wetterbedingungen sehr gut zurechtgekommen! Hut ab!*[101]
'First I drove into the mountains with these tires, where, unexpectedly, **snow was snowing**. The tires coped very well with these weather conditions! Hats off!'

[101] Source: www.oponeo.at/reifen-modelle/imperial-ecodriver-3 oponeo/tires-models-imperial-ecodriver-3 (last accessed March 15, 2024)

There is no definite article since there is no mention of any precipitate in the preceding context. An indefinite article would have been possible but odd, especially without any further modification. Here, too, it would have been unproblematic to use the predicate construction (*wo es unerwartet schneite* 'where it was unexpectedly snowing') or an argument construction (*wo unerwartet Schnee fiel* 'where, unexpectedly, snow was falling'). As opposed to (226) and (227), there is no apparent motivation to employ the cognate construction, as the focus of the message is the performance of the car tires (in an unexpected weather condition) rather than the construal of a particular mood or atmosphere.

In example (230), taken from the weekly German newspaper supplement *Süddeutsche Magazin*, the rain is introduced in the first sentence with *ein Tag um den anderen Regen* 'one day after the other rain' in a predicateless construction, where an argument construction, [*es gibt Regen*] or [*Regen fällt*], would appear to be fitting or even a preposition such as *mit* 'with' (*ein Sommer … mit Regen* 'a summer with rain').

(230) *Es war die längste Zeit ein erbärmlicher Sommer gewesen, ein Tag um den anderen Regen. Die Boulevardblätter hatten sogar damit begonnen, ihn auf ihren Titelseiten anzuklagen, wie sie sonst nur das S-Bahn-Chaos anklagten, aber **der Regen regnete** einfach **weiter**.* (sz-magazin.sueddeutsche.de 'South German Newspaper Magazin')
'It had been a miserable summer for the longest time, one day after the other rain. The tabloids had even started denouncing it (author's note: lit. 'him', the rain) on their front pages, as they usually only denounced the subway chaos, but **the rain** just **kept raining.**'

In the sentence that contains the cognate construction, the rain is then represented as some sort of personified, volitional entity, which, despite being denounced by the tabloids, continues with what it is doing without care or consideration. The particle *einfach* 'simply/just' serves to express either a sort of perplexity at "the rain's brazenness" or gives the utterance a more emotional tinge. Take note that the verb in the cognate construction occurs in combination with the separable prefix *weiter-* 'further/continue…/keep', also an independent adverb, which marks the construction as imperfective/ongoing.[102]

The WEATHER TALK examples of the cognate construction thus show a mix of ordinary day-to-day use as would occur in spontaneous spoken discourse, as in (225) and (228), and slightly more poetic, or at least more elaborate written discourse, where the more carefully thought-out

[102] *Es soll weiterregnen* 'it is supposed to continue raining'; other examples of verbs with prefix *weiter* are *weiterleben* 'continue to exist', *weitermachen* 'proceed, keep going', *weitersprechen* 'keep/continue talking', etc.

wording serves to construe a specific atmosphere, as in (226) and (227). The former could easily be substituted with the "default" predicate construction. In fact, we might expect the predicate construction over the cognate construction in these instances. Take note that there is little, if any, additional modification and the utterances occur mostly in combination with a definite article, but not exclusively so, as in (229).

Research has shown, in fact, that one factor that governs the acceptability of cognate constructions and, thus, the acceptability of precipitates as cognate objects is extra-propositional content (e.g., Austin 1982, Horita 1996, Rice 1987). Rice (1988: 207–210) found, for English, that some sort of "special modification", such as "a definite or demonstrative article, an adjective, or a relative clause" (ibid.: 208), which sets up the cognate object as a distinct entity is the main factor, which controls the permissibility of cognate objects. The modification helps the speaker to delineate the process from the processed object, as illustrated in (231). The examples are reproduced from Rice (1988: 208).

(231) a.**Susan lived a life.*
b. *Susan lived a good life.*
c. *Susan lived the life she wanted.*
d. *Susan lives her life well.*

Similar observations have been echoed, among others, by Horita (1996), who found that, In English, cognate object constructions can be replaced by intransitive constructions containing an adverb where the construction has a generally synonymous reading, e.g. *She laughed a raucous laugh* and *She laughed raucously* (Horita 1996: 221–222). Interestingly, none of the examples from (225)–(230) contain modification in a way where the very instance of the ME is clearly distinct from other instantiations. In (225)–(228), the only modifier is a definite article, but the examples appear just as acceptable if the definite article were left out. Example (228) is the only one in which the ME is introduced in the (returned concordance) context directly preceding the sentence with the cognate construction.

Similarly, in (232) and (233), *Schnee* 'snow' is set apart as a distinct entity by defining it as *a very wet snow, which will freeze* in the former and as *mushy snow* in the latter. These two examples differ in that (232) specifies the characteristics of the snow on the very day the utterance is made (*heute* 'today'), while (233) specifies a particular type of snow that presumably is typical of a certain non-specified region (*da wo* 'there where').

(232) *Wird eine Überraschung, da **es heute einen sehr nassen Schnee schneite** und der nun gefrieren wird dank Nord-Ost-Wind und den entsprechenden Temperaturen.*
(tumbral.com/blog/diefrankfurterin 'blog/the frankfurter')
'Gonna be a surprise, as **it snowed a very wet snow today** and it will now freeze thanks to the north-east wind and the corresponding temperatures.'

(233) *Da wo es immer grau und trist ist, regnet oder **matschigen Schnee schneit**, würde ich das auch tun.* (*DWDS* – regenbogenlichter.wordpress.com 'rainbowlights.wordpress')
'There where it is always grey and drab, rains or **mushy snow snows**, I would do the same.'

The second most obvious type of cognate construction does not contain much modification either except for some definite articles: these are the highly marked utterances from poems and song lyrics, which in Chapter 3 would have fallen under the LITERARY genre category. One of these instances is in fact the earlier mentioned example provided by the *DWDS* dictionary: *der Regen regnet jeglichen Tag* 'the rain it raineth every day'. In line with this example, a total of seven concordance lines returned from the *deTenTen13* either quote Nicolai's opera *Die lustigen Weiber von Windsor The Merry Wives of Windsor* or refer to Shakespeare's *Twelfth's Night* and the *Fool's Closing Song*, which is the origin of the expression *the rain, it raineth every day* and, correspondingly, Nicolai's adaptation.[103] The expression is often interpreted as a metaphor—more specifically as a "note of quotidian sadness"—which serves as a reminder of the fact that life is difficult and full of hardships (*wind and rain*) where merriment and celebrations are merely fleeting interludes in the daily grind.[104] The *DWDS* corpus returns did not include this particular expression, but a very similar one—or rather two––by the French-German poet, writer, and botanist, Adelbert von Chamisso. In his ballad, *Der Szekler Landtag* 'The Szekler State Parliament' (dated 1831), Chamisso describes a worrying situation where a ceaseless rain threatens to ruin the farmers' crops in *Siebenbürgen* 'Transylvania'. The lines in question are given in (234) and (235).

(234) ***Es regnete der Regen** alle Tage* (line 8)[105]
'it rained the rain every day'

(235) *Ihr seht, **der Regen regnet** ewig fort* (line 39)
'you see, the rain is raining forever'

[103] *Feste's Song* from *Twelfth Night* by Shakespeare, titled *The Fool's Closing Song.*

[104] Source: shakespearetheatre.org (last accessed February 5, 2024)

[105] Source: www.lyrikkompass.de/poems/56353 (last accessed February 20, 2024) for the complete poem with line numbering.

While the ballad is in rhyme form, it is not overly lyrical or poetic. It deals with a common, tangible problem, which, and this is the crux of the matter, cannot be solved by human intervention. In this case, incessant rain has been falling for weeks and is threatening to damage the crops. The people of Siebenbürgen appoint a state parliament for the purpose of solving the matter, but the state parliament keeps postponing finding a solution until they ask one of their elders for advice. The elder suggests watching the situation for another 14 days and if the rain has not stopped raining by then, well, then it should continue to rain "as long as it wants and can". In the end, the sun starts to shine and dries the fields, and an abundant harvest is reaped. As such, the ballad apparently is meant as a plea for sanity and the advice that time will fix things. It is an allegory, but it can obviously also be understood literally. Example (234) occurs a total of 36 times in the *DWDS* metacorpus *Historische Korpora* 'Historical corpora', which contains texts from the years 1465–1969, and example (235) occurs six times more (N=42). The metadata reveal that the ballad was printed in many different *Lesebücher* 'reader/reading book' for pupils. Thirty-eight results contain the word *Lesebuch* 'reader/reading book' in the metadata of their entry. Eight of those Chamisso quotes belong to different reading books published for different schools (e.g., boys' only and girls' only schools in different regions and for different grades) but have the same editors. Thus, it may not be surprising that the readers contain similar materials.

Nicolai's opera (1845–1849) was composed not much later than Chamisso's ballad but does not feature by far as frequently as these two lines by Chamisso. Still, Chamisso and Nicolai's utterances are the most frequently repeated items overall, and both contain cognate ME constructions. There are other returns from the LITERARY genre, although not quite as many as anticipated. The *deTenTen13* contained only two more examples from the LITERARY genre. Since the *deTenTen13* is a web corpus, I expected the returns to contain more utterances from poems, song lyrics, and LITERARY sources in general. One of these two LITERARY examples is from the German singer-songwriter, poet, and former East German dissident, Wolf Biermann, and it was also returned in a *DWDS* corpus search. The other example is from the more contemporary pop group *Juli*. Biermann's use of the cognate construction differs from other lyrical uses in that it is vividly descriptive but not metaphorical. The song is called *Ballade vom Traum* 'Ballad of the Dream', although it appears to depict a nightmare, rather than a dream, of an unsuccessful attempt to escape from the former GDR. In the song, Biermann describes thrashing along the (Berlin) wall and tearing through the barbed wire when his stomach is ripped open by (gun)

shots, followed by his guts pouring out and being devoured by a German shepherd. Biermann was expatriated from the GDR in 1976 and the ballad is rather about the impossibility of himself returning into the GDR. However, the actual likelihood of being shot by the GDR border troops obviously concerned GDR citizens trying to escape into West Berlin more so than Western Germans trying to "break into" the GDR. The lyrics then continue as in (236).

(236) *Roter Saft fällt komisch aus der tauben Haut.*
*Und **Regen regnet** in den starren himmeloffnen Mund*
– das ist mein Traum, vor dem mir jeden Abend graut. (zeit.de 'Times online newspaper')

'Red juice falls strangely from the numb skin
And **rain rains** into the rigid, sky-open mouth
– that is my dream, which I dread every night.'

Although the *deTenTen13* returned fewer instances of LITERARY utterances than expected, both the *DWDS* corpora as well as a Google search produced many more such instances. Undoubtedly, this was due to the fact that, on the internet, there are numerous platforms dedicated to the collection and discussion of song lyrics as well as a wide access to poetry and other literary forms. Example (237), likewise from the *DWDS* historical corpus, is LITERARY in that it comes from a late 19th century narration, but like many of the literary examples it does not differ much, if at all, from the weather talk utterances, as in (226) and (227), which concern actualized MEs.

(237) *Ich (…) legte den Arm (…) um sie; und sie hatte den Kopf an meine Schulter gelegt, und **der Regen regnete** immer zu, und wir ließen ihn glückselig dabei.* (*DWDS* – Wilhelm Raabe, 1884. *Pfisters Mühle: Ein Sommerferienheft* 'Pfister's Mill: Notes From a Summer Vacation')
'I put my arm around her; and she had put her head on my shoulder, and **the rain was raining** incessantly, and we blissfully let it.'

Example (238), from a poem by the 19th century German poet and painter, Max Dauthendey, contains both precipitates, *Regen* 'rain' and *Schnee* 'snow, in their respective cognate constructions with temporal modification.

(238) *Noch **schneit es Schnee** in einer Stund', Und **regnet Regen** in der andern; Der Frühling ist in aller Mund.* (wortgewimmel.de 'word swarm', projekt-gutenberg.org)
'It still **snows snow** one hour, and **rains rain** the next; Spring is on everyone's lips.'

The topic of the poem is the transition from winter to spring and is related by a description of the fickleness of the weather where it is still cold enough to snow in one hour and the next it is already warm enough for the precipitation to fall as rain. Example (239), from a *Google* search,

involves another simple ME description in rhyme form. What makes it interesting is that the poem was written by a fifth grader, presumably 10–11 years old.

(239) ***D'r Schnee schneit so schee*** **'The snow snows so nicely**
alle hocket in d'r Stub all are sitting in the parlor
und trinket Tee and drink tea'

(*Verrückte Weihnachten* 'Crazy Christmas' by Laura, 5th Grade HS Diepoldshofen)[106]

She, too, modifies the snow event with the manner adverb *so schee*, which is a regional variant of Standard German *so schön* 'so nicely/beautifully'. Obviously, the first line also contains a nice triple alliteration. Like the Biermann lyrics, the constructions here refer to actual weather events, while the lyrics in (240) may be metaphorical. Here, the singer-songwriter, Götz Alsmann, expresses how he is able to brave adverse weather conditions (snow, wind, storm), because the love of another person is keeping him warm.

(240) ***Der Schnee schneit heftig,*** **'The snow is snowing fiercely**,
Der Wind bläst kräftig, The wind is blowing hard,
Doch ich trotz dem Sturm mit dir im Arm. But I brave the storm with you in my arms.'
(*Denn meine Liebe hält mich warm* 'For my love keeps me warm' by Götz Alsmann)[107]

The snow event is modified by the adverb *heftig* 'severely', which appears to be a frequent modifier of both *Regen* 'rain' and *Schnee* 'snow' in cognate constructions. While the adverse weather conditions could of course be metaphorical for adverse life events, the lyrics read like the literal description of a person experiencing cold weather and may very well be understood literally as the song is furthermore part of a "winter album" called *Winterwunderwelt* 'winter wonder world'.

Some of these concordance lines, whether from the LITERARY genre or not, in a similar manner to Chamisso's, Nicolai's and/or Shakespeare's line, express the fact of something unavoidable that is happening on a regular basis, cannot be stopped, and is not within the control of the speaker. This is sometimes emphasized by the addition of the verb *lassen* 'let/leave', where the speaker gives the advice to 'let the rain rain' (*lass den Regen regnen*): it is

[106] Source: docplayer.org/26966349-Wintergedichte-der-klasse-5-hs-diepoldshofen.html - 'winter-poems-of-grad-five-secondary-school-Diepoldshofen' (last accessed April 07, 2024)

[107] Source: www.musixmatch.com (last accessed April 07, 2024)

what it does and there is no use in fighting it or complaining about it as in (241). This example is taken from song lyrics; it is the chorus of a song called *Tränen* 'tears' by artist *Haller*.

(241) ***Lass den Regen regnen****, lass die Sonne scheinen,*
Deine Tränen trocknen schon von allein. (*Google*)
'**Let the rain rain**, let the sun shine,
Your tears dry on their own.'

The singer sings about happily providing a shoulder to cry on for a friend and advising the friend to let things run their course and everything will turn out fine in the end. The expression is put in juxtaposition with the more positive ME *lass die Sonne scheinen* 'let the sun shine'. While the reference to the MEs may be metaphorical for "let the bad and the good happen to you", it is also feasible that the singer intended to refer to the actual MEs as in "time passing" and rainy and sunny days needing to pass for healing to happen. By contrast, the ME in (242), from a travel blog, is actualized.

(242) *Wir erledigen noch einigen Kleinkram,* ***lassen den Regen regnen****, bleiben im Trockenen und hoffen, dass das Wetter morgen doch etwas besser sein wird*.[108] (*Google*)
'We still run a few errands, **let the rain rain**, stay dry and hope that the weather will be a little better tomorrow.'

The writers of the blog travel by boat and report running errands during a dry phase of the day. While it is raining, they stay aboard their boat *im Trockenen* 'in the dry' and let the rain "do its thing" and hope for better weather for the next day. Obviously, the meaning of *lassen* in these two examples is different from causal *lassen* 'let/cause', which we find in combination with agents like a god or God in the scripture, as in (243). These examples involving transitive, causal agentives are further discussed in Section 4.3.1.

(243) *Der HERR ist es, der die Wetterwolken macht,* ***er lässt den Regen regnen*** *(…)*
(wort-und-wissen.de 'word-and-knowledge')
Lit. 'It is the LORD who makes the weather clouds, **he lets the rain rain**'

One type of cognate construction that appears more frequently for the *Schnee* 'snow' precipitate (N=6 in the *deTenTen13*) than for the *Regen* 'rain' precipitate (N=0 in the *deTenTen13*) is a resultative one. The reported volumes of snow are gauged and reported in centimeters, just like rain is gauged and reported in millimeters, and where one centimeter of

[108] Source: my-madness.ch/letzter-tag-in-eckernfoerde-2 'my-madness/last-day-in-Eckernförde' (last accessed March 15, 2024)

snow is equivalent to about 1 millimeter of water once the snow is melted. Examples (244)–(247) are all informal WEATHER TALK situations. The first two examples, (244) and (245), measure the snowfall in centimeters. The quantities in (246) and (247) are "less formal" and on opposite sides of the spectrum: the former de-emphasizes the quantity as *3 Flöcken* **'**three (little) flakes' and the latter refers to the unspecific *Unmengen* 'oodles'. *Mengen* translates to 'quantities/masses/lots'. The negative form *Unmengen* is a more emotionally charged intensifier basically meaning the same, i.e., great quantities or rather an excess quantity of something. The emotional declaration continues with the interjection *mein Gott* 'my God' and announcing that the speaker fears their commute the following day due to the danger of black ice since it started raining in the meantime. In (247), there are two different constructions given: first, the cognate construction in combination with the *Schnee* 'snow' precipitate and then the predicate construction in combination with the *Regen* 'rain' precipitate. The cognate construction appears to be perfective while the *es*+PRED-Cx is happening concurrently to the message being written as indicated by the adverb *nun* 'now'. Aside from (244), which is "bare" except for the quantity, the other three examples, furthermore, include qualitative modifiers and stancetaking in addition to the quantitative modifier. While the quantifier suffices to license the cognate type, any additional modification, such as MANNER or STANCE, appears to make the construction even more acceptable (cf. Rice 1988).

(244) *Es ist nicht leicht, an den Sommerurlaub zu denken, wenn **es** **20 cm** **Schnee** **geschneit** hat.*
(alltagsschrott.de – 'everydayjunk')
'It isn't easy to think of summer vacation when **it** **has** **snowed** **20 cm of** **snow**.'

(245) *Hier **hat** **es** heute wieder **10 cm wunderbar weißen, wattigen** **Schnee** **geschneit**!*
(daisysdasein.blogspot.ch – 'Daisy's-existence.blogspot)
'Today **it** **snowed** again **10 cm of wonderfully white, fluffy** **snow**!'

(246) *@Wilde Henne: schon wieder Sommer, neinnein, die Uhr halten wir lieber an, jetzt, wo **es** endlich **3 Flöcken** **Schnee** **geschneit** **hat**.* (*DWDS*, lamiacucina.blog)
'@Wilde Henne: Summer again, nono, we rather stop the clock, now that **it** **has** finally **snowed** **3 (little) flakes of** **snow**.'

(247) *Bei uns **hat** **es** heute **Unmengen** **Schnee** **geschneit**, nun regnet es, mein Gott, ich muß morgen zur Arbeit, fürchte mich jetzt schon. Glatteisgefahr!!*
(seelensachen.blogspot.ch – 'soulthings.blogspot')
'It **snowed** **tons of** **snow** here today, now it's raining, my God, I have to go to work tomorrow, I'm frightened already. Black ice danger!!'

Examples similar to the ones above for the *Schnee* 'snow' precipitate can easily be found on Google as well as in the *DWDS* corpora. Interestingly, the cognate construction does not

appear to be used resultatively for the *Regen* 'rain' precipitate although rainfall is clearly measured in a similar manner.

With a pluviometer or rain gauge, the amount of rain that has fallen is measured in the same way as snowfall is measured. Yet, it is difficult to come by resultative examples which state, for example, that *it rained 20mm of rain*. Instead, technical weather reports appear to default to either the argument construction, [*es geben*] or [*fallen*], where the *Regen* 'rain' precipitate is concerned. Obviously, the visual of a snowfall is more salient than that of a rainfall and can be "measured" somewhat reliably by eyeballing the precipitate on the ground, which is presumably the reason for speakers to employ the resultative construction with the Schnee 'snow' precipitate.

With that being said, the resultative cognate construction with the *Schnee* 'snow' precipitate is not the principal choice in weather forecasting either. Still, examples of this kind can more easily be found than for the *Regen* 'rain' precipitate. Generally, the resultative cognate types are the only ones, which additionally contain the pronoun *es* 'it' in subject position. This makes them appear to have a higher level of transitivity as the setting/participant amalgamation *es* 'it' acts as an agent (*es hat 20cm geschneit* 'it snowed 20cm'). It is not possible to make *Schnee* 'snow' the subject and/or causal agent: **Schnee schneite 20cm* 'snow snowed 20cm'.

Lastly, there are two resultative cognate construction examples involving *Hagel* 'hail', or rather, the object is *Hagelkörner* 'hail stones', as in (248) and (249). In both cases, the object is additonally modified in a qualitative manner with the size of the hail given as *größer* 'larger' and *zentimetergroß* 'centimeter-sized'. In (248), the hail stones are further modified by a relative clause (*that will damage the car*), which suggests that the size of the hailstones will do damage. These two examples are, however, not the returns listed in Table 98, as the search query was carried out for the "word" *Hagel* 'hail' and not the lemma, which did not include compounds like *Hagelkörner* 'hailstones'.

(248) *Die Gefahr, daß **es größere Hagelkörner hagelt**, **die das Auto kaputt machen**, ist auch gegeben*. (finanz-gruppen.de 'finance-groups')
'There is also a risk that **it hails larger hailstones that will damage the car**.'

(249) ***Es hagelte** ab 15:30 Uhr zwischen Hofolding und Aying ca. 10 Minuten lang **zentimetergrosse Hagelkörner**, und es bildete sich eine Hagelschicht.* (sturmwetter.de 'stormweather')
'From 3:30 p.m. between Hofolding and Aying, **it hailed centimeter-sized hailstones** for about 10 minutes, and a layer of hail formed.'

Finally, there are the tautological cognate constructions which are necessarily true and apparently uninformative. These utterances are usually not unmotivated, however. In (250), a LITERARY example, the sentence preceding the ME spells out that everything has its purpose or just a *raison d'être*, which for *rain* is *to rain*. It just so happens that a weather verb's subject or object tends to be etymologically related to the verb – in fact they are the very same thing.

(250) *Sie erklären ihm, dass jeder von ihnen eine eigene Aufgabe erfüllt oder einen eigenen Willen hat.* ***Der Regen regnet****, die Katze sonnt sich und der Schatten erinnert daran, dass alles zwei Seiten hat.* (litrix.de 'German literature online')
'They explain to him that they each fulfill a task, or have a will of their own. **The rain rains**, the cat basks in the sun, and the shade reminds us that everything has two sides.'

While the speaker in (251) clearly is bored by the foreseeability of what the precipitates do (*Schnee schneit, Regen regnet, Hagel hagelt*), the speakers in (252) and (253) give information about the current weather situation, but they do so by playing with the cognate concept and applying it to other items as well: *die Kälte kältet* 'the cold colds', *die Rolle rollt* 'the roll rolls', and *Nacht nachtet* 'the night nights'. The ambient and diurnal cognate verbs sound odd, but they are indeed items in the *Duden*: *nachten* 'to night' means 'to become night, get dark' and *kälten* 'to cold' means 'to become cold' or 'cause someone to feel cool due to its nature'. *Nachten* 'to night' is apparently frequently used in Swiss German and rare and poetic in Standard German. *Kälten* 'to cold', too, is rare. These extra cognate constructions are also licensed by the list-like exposition.

(251) *Es schneite also, ich glotzte hinaus und fand es einfallslos von der Schöpfung, dass immer nur Flocken vom Himmel fallen oder Tropfen, eins von beiden, vom Hagel als Ausnahme mal abgesehen.* ***Schnee schneit, Regen regnet, Hagel hagelt****.* (sz-magazin.de)
'So it was snowing, I stared outside and found it unimaginative of creation that only ever flakes or drops fall from the sky, one of the two, with the exception of hail. **Snow snows, rain rains, hail hails**.'

(252) *Denn das Wetter ist alles andere als gut. Miserabel trifft es wohl eher. Es ist kalt, uselig und kabbelig.* ***Der Wind pfeift****,* ***der Regen regnet****, und* ***das Kalt kältet***. (transitfrei.de 'transit free')
'Because the weather is anything but good. Miserable is more fitting. It's cold, uncomfortable and choppy. **The wind whistles**, **the rain rains**, and **the cold colds**.'

(253) *Peter hat den Blog im tiefsten Winter eröffnet:* ***Der Schnee schneit****,* ***die Rolle rollt****,* ***die Nacht nachtet*** *und es sind -10°C draußen.* (*DWDS,* 'teamlobre.wordpress.com')
'Peter started the blog in the dead of winter: **the snow snows**, **the roll rolls, the night nights** and it is -10°C outside.'

Since *kälten* 'to cold' is rare and ?*das Kalt* is odd for what, should be the more grammatically correct *die Kälte* 'the cold' (which would be an even better cognate because of

the matching *umlaut*), it is surprising that the speaker in (252) did not choose a cognate construction instead of *der Wind pfeift* 'the wind whistles', such as *der Wind windet* 'the wind winds', since *winden* is, in fact, an acceptable verb that is rare as well, but, nevertheless, part of the *Duden*. While *wind* does not always whistle, the collocation could possibly be frequent enough to count as a *split construction* where both elements, the predicate and the argument, encode the ME. More appropriate split constructions are *der Wind weht* 'the wind blows' and *der Wind bläst* 'the wind blows', but since this dissertation treats, first and foremost, precipitation events, the next section concerns the split construction in the German language, and more specifically the verb *rieseln* 'trickle/drift/flutter' in combination with the *Schnee* 'snow' precipitate and to a lesser degree with the *Regen* 'rain' precipitate.

4.2.2 The split type

The second predicate-argument type of ME construction being considered here, which divides the semantic load between the two components, is the split type. Here, the ME is encoded in either the argument or the predicate while the respective other component encodes some other facet of the ME. This is similar to expressions examined in the last section with *Wind weht* or *Wind bläst* 'wind blows', where *wehen* is something only the wind can do (or things like laundry, hair, or flags can *wehen* 'blow/flutter' **in** the wind). Both verbs can also be used in a predicate construction: *es bläst ganz schön draußen* (cf. *Duden*, 'it's blowing quite a bit outside') and *es weht* 'it blows'. Since this dissertation is mostly concerned with precipitation events, this section focuses on the verb *rieseln*, which occurs with a number of noun collocates, but with regard to MEs frequently collocates with the *Schnee* 'snow' precipitate and, to a lesser extent, with *Regen* 'rain'.

4.2.2.1 Frequencies and usage of the precipitate verb rieseln *'flutter/trickle/drift'*[109]

As with the cognate type, the split type is not the predominant construction choice in German, but instances of the split type within the precipitation domain immediately come to mind. Other than introspection, one way to determine which verbs frequently collocate with the various

[109] The translations for *rieseln* are sourced from the online dictionary *dict.cc*.

precipitate types is to examine and survey SE's *Word Sketch* function, which sheds light on both the most frequent verb collocates as well as the strongest verb collocations. As a reminder, this is done by means of a *logDice* score where "comparing two scores, plus 1 point means twice as often collocation, plus 7 points means roughly 100 times frequent collocation [sic]" (Rychlý 2008: 9).[110] For the *Schnee* 'snow' precipitate, the fifth-highest verb collocate is *rieseln* 'flutter/trickle/drift' (N= 2,560). Since there is no direct equivalent for *rieseln* in English, I will alternate between the different possible translations depending on the context even if the translation seems odd (e.g., 'flutter'). I will also resort to simply using 'fall' in the gloss. In turn, the *Word Sketch* for verb *rieseln* returns *Schnee* 'snow' as its top subject collocate by a wide margin. Despite "only" being the fifth verb collocate of *Schnee* 'snow, *rieseln* has the highest *logDice* score out of all verb collocates (10.1), which makes it the strongest collocate to occur with the noun *Schnee* 'snow'. The "theoretical maximum" of the *logDice* score is 14. This applies to cases "when all occurrences of X co-occur with Y and all occurrences of Y co-occur with X. Usually the value is less than 10." (Rychlý 2008: 9). In comparison, the collocating verb *fallen* 'fall' has an almost three times higher incidence (N= 7,512), but only a *logDice* score of 7.2. The verb collocate *kommen* 'come' has the fourth highest incidence (N=3,157), but a low *logDice* score of 3.6.

In addition to providing some frequencies, this section explores the semantics and usage of the verb *rieseln* 'flutter/trickle/drift' (as well as its *anlaut*-variant *nieseln* 'drizzle') as one of the predominant collocates with the *Schnee* 'snow' precipitate, which means that the ME is encoded in the argument. Since *rieseln* does not exclusively occur with *Schnee* 'snow', other subjects are explored as well, not only based on the *deTenTen13* corpus, but also with respect to the German corpus *DWDS.* The *DWDS* was used to cross-check the *deTenTen13* results since the latter mostly contains more contemporary utterances. By contrast, the *DWDS Wortprofil* 'word profile', which is similar in its interface to the *SE Word Sketch* function, is based on a cross-section of the *DWDS* corpora and includes historical data.[111] The *Schnee+rieseln* 'snow+flutter' collocation is interesting insofar as the verb is far from "light" (as opposed to *fallen* 'fall' or *kommen* 'come') and chiefly occurs with the *Schnee* 'snow' precipitate. In the past, *rieseln*

[110] The sentence is a direct quote including any and all grammatical errors.

[111] Source:www.dwds.de/b/dwds-wortprofil-in-neuer-version/ 'word profile new version' (last accessed February 29, 2024)

seems to have occurred more frequently with the precipitate *Regen* 'rain'—a development, which will be explored later in this section. As a general rule, *rieseln* hardly collocates with other types of precipitates or MEs in more contemporary German and does not appear in the *Word Sketch* top five, or even top ten, verb collocates of other precipitates. The following paragraphs list various definitions, the etymology, and a number of examples for *rieseln* from the *Duden* I, the *DWDS* dictionary (II) and the brothers Grimm's dictionary a(III), the *Deutsches Wörterbuch* 'German dictionary' (henceforth *DWB*), starting with the current definition in the *Duden*, which is regularly updated:

I. Definition of *rieseln* (*Duden*)
   Mit feinem, hellem, gleichmäßigem Geräusch in vielen kleinen Teilchen leise, kaum hörbar nach unten fallen, gleiten, sinken[112]
   'Fall down hardly audibly, glide, sink; with fine, clear sound, steadily, in many little particles'
II. Definition and etymology of *rieseln* (*DWDS*)
   The definition for *rieseln* is given in (a)–(b) and the etymology is given in (c).[113]

   a. *spärlich fließen, in feinen Tropfen, unzähligen Partikeln herunterfallen*
      'flow sparsely, falling down in fine drops, countless particles'
   b. *kaum hörbar fließen, in feinen Tropfen, unzähligen Partikeln fast geräuschlos herunterfallen* 'flow barely audible, falling down almost silently in fine drops and countless particles'
   c. *sacht fließen, rinnen*, *in feinen Tropfen niederfallen* 'flow softly, trickle, fall down in fine droplets' from Middle High German (dated between 1050 and 1350) *riselen*: *tröpfeln*, *leise regnen* 'drip, rain softly', iterative

The definitions in (I) and (II) demonstrate how *rieseln* not only implies the manner of movement, but also suggests a specific sound that may accompany the movement of the FIGURE or, rather, the mass of the FIGURE, as made up of many small components. We can hardly claim, then, that a verb like *rieseln* is semantically "vacuous" or even just "light". In fact, as we have seen, the semantics of *rieseln* cannot be homogeneously translated into a single English verb, but, depending on the subject, translate to 'flutter', 'trickle', or 'drift'. With respect to the PATH of the FIGURE, the current *Duden* definition states *nach unten (fallen)* '(fall) down(wards)', although regarding possible subject collocates such as *water* in a stream or a creek trickling over rocks or

[112] Source: www.duden.de/rechtschreibung/rieseln (last accessed February 15, 2024).

[113] Source: www.dwds.de/wb/rieseln (last accessed February 15, 2024). Additionally, the etymology states that *rieseln* is an iterative form of the Old High German verb *rīsan,* meaning to 'fall, fall down' but also Middle High German *rîsen* 'to mount, fall' according to en.wikisource.org/wiki/An_Etymological_Dictionary_of_the_German_Language/Annotated/Reise (last accessed April 1, 2024).

gravel, the direction may arguably be sideways or a combination of down plus sideways (see fn.112). With respect to *Wasser* 'water', it appears that the sound component of the definition prevails over the PATH component. Lastly, the *DWB* by Jacob and Wilhelm Grimm, arguably the largest and most comprehensive dictionary of the German language in existence, provides a range of historical definitions and examples for *rieseln*, spreading over a couple of pages, with the ones of importance to this dissertation given in (III). Here, the definition as well as the examples by different linguists (their names are given in small caps) provide a range of precipitates, such as rain, snow, hail, and fog as well as the path (*herunter* 'downwards'), the MANNER (e.g., *sanft* 'softly'), and the QUANTITY (e.g., *ein wenig* 'a little').[114]

(III) Definition *rieseln* (*DWB*)

sanft herunterfallen 'fall softly' and

a) *in einzelnen theilchen, tropfen- oder körnerweise, so*
'in individual particles, drop- or grainwise, so'

1. *zunächst von regen- und thautropfen, schneeflocken und hagelkörnern* [sic]
'to begin with of rain- and dewdrops, snowflakes and hailstones'

- *…unpersönlich, riselen, ein wenig regnen, rorare* SCHOTTEL
'impersonal, *riselen*, rain a little, be moist/cause dew/drop';[115]
- *risle, hageln* HUNZIKER '*risle,* hail';
- *rieseln*, *minutim grandinare* STIELER '*rieseln*, hail in little pieces/minutely';
- *risseln, rieseln, fein regnen* WOESTE '*risseln, rieseln,* rain finely';
- *riseln, fein regnen; wenn ein starker nebel seine tropfen niederschlägt…* SCHAMBACH '*riseln*, rain finely, when a heavy mist precipitates its drops';
- es rieselt *sagt man auch, wenn der gefrorne schnee in gestalt kleiner hagelkörner einzeln herunter fällt. an andern orten* rieselt es*, wenn es in einzelnen kleinen tropfen regnet.* ADELUNG
'they also say that *es rieselt* when the frozen snow falls down individually in the form of small hailstones. In other places *rieselt es* when it rains in small drops'*;*
- *als wenn es linde schnee-flocken gerisselt hätte …* 'as if it had trickled soft snow-flakes'

b) *sanft, mit leichtem gefälle flieszen; so vom wasser, namentlich dem eines rinnsals, einer quelle, eines baches: rieselen, sanft flieszen, fluescere* (…) [sic]
'flow gently, with a slight descent; so of water, namely that of a trickle, a spring, a stream: trickle, flow gently, *fluescere*'

---

[114] The spelling in (a(III)) is archaic. The translations in (I)–(a(III)) are provided by the author. The examples range approximately from the mid 17th century to the mid 19th century. The *Duden*, the *de facto* official dictionary of the Standard High German language, was first published in 1880, so that the spelling of *rieseln* varies and can be found, for example, as *rieseln*, *riseln* and *risseln*. Source: woerterbuchnetz.de/?sigle=DWB&lemid=R05404 (last accessed February 15, 2024).

[115] Source: Latin-dictionary.net (last accessed April 01, 2024).

The different definitions place the focus on different aspects of the verb. While the *Duden* definition does not specify the nature or quality structure of the argument or FIGURE at all (other than specifying that it is made up of many small particles), Grimm's definition lists the types of precipitation as well as "types of water", among others. In Grimm's definition, several precipitates are provided, with none given preference over another, but the historical examples seem to give preference to the rain and hail precipitates. It appears then, that the predominance of the *rieseln*+*Schnee* 'flutter+snow' collocation, as established by the *Word Sketch*, has tightened over the years. For a start, this claim is based on *rieseln* being the top fifth verb to collocate with *Schnee* 'snow' in the *deTenTen13*, while it does not show up at all in the verb collocates of precipitates *Regen* 'rain', *Schneeregen* 'sleet', or *Hagel* 'hail'. Secondly, one of the top "subject" collocates for *rieseln* in both the *deTenTen13* and the *DWDS* is the adverb *leise* 'softly/quietly'. Presumably, it is wrongly listed as "subject" due to its sentence-initial position in the multi-word sequence *Leise rieselt der Schnee* 'Softly falls the snow', which probably can be ascribed to the song of the same name as the multi-word sequence, first published in 1895. *Leise rieselt der Schnee* is one of the most popular German winter or advent season songs, which the author himself, Protestant pastor Eduard Ebel, called a "children's song" (cf. Mang and Mang 2007: 1005–1006). There are more than 1,000 records and CDs with the song in the *Deutsches Musikarchiv* 'The German Music Archive', which is part of the *The German National Library* in Leipzig. The only winter song to appear on more storage media is *Stille Nacht, heilige Nacht* 'Silent Night' (roughly 2,000 records and CDs).[116] A *Google* search for the title in combination with "YouTube" returns almost 300 videos (on *Google's* "Videos" tab and not "All") from *YouTube* just for the song alone, not including any compilations, etc. A basic search query for *leise rieseln der Schnee* in the *deTenTen13* returns 1,722 concordance lines and *leise rieseln* returns 2,291 concordance lines. The search queries were carried out for the lemma *rieseln* in both the *deTenTen13* and the *DWDS*. Furthermore, *DWDS* returned 1,374 concordance lines for the search query *leise && rieseln && Schnee*, where the word order is not fixed. In the *DWDS Webkorpus*, a total of 1,162 instances are exact matches of *leise rieselt der Schnee*.

This assumption seems to be confirmed by querying *Google Ngrams*, as given in Figure 43.[117] The *Google N-gram* viewer is a tool that displays a frequency graph showing how a

[116] Source: deutschelieder.wordpress.com/tag/kinderlied/ 'German songs/children's songs' (last accessed April 10, 2024).

[117] Source: Google Books Ngram Viewer books.google.com/ngrams (last accessed March 7, 2024).

queried word or phrase has occurred in a corpus of books over time. For the 4-gram *Leise rieselt der Schnee*, the diagram shows a flat line up until approximately 1920 and then a steady increase up to the 2000s. The time range 1800–1920 shows a sharp increase between 1905 and 1914, which has been embedded in Figure 43. It is feasible that the song gained popularity in that time period and afterwards the expression was distributed more widely. There are no zero instances of the 4-gram returned by *Google Ngrams* for the years 1800–1905. Obviously, there are more books published nowadays, but the results are normalized by the number of books published in each year.[118]

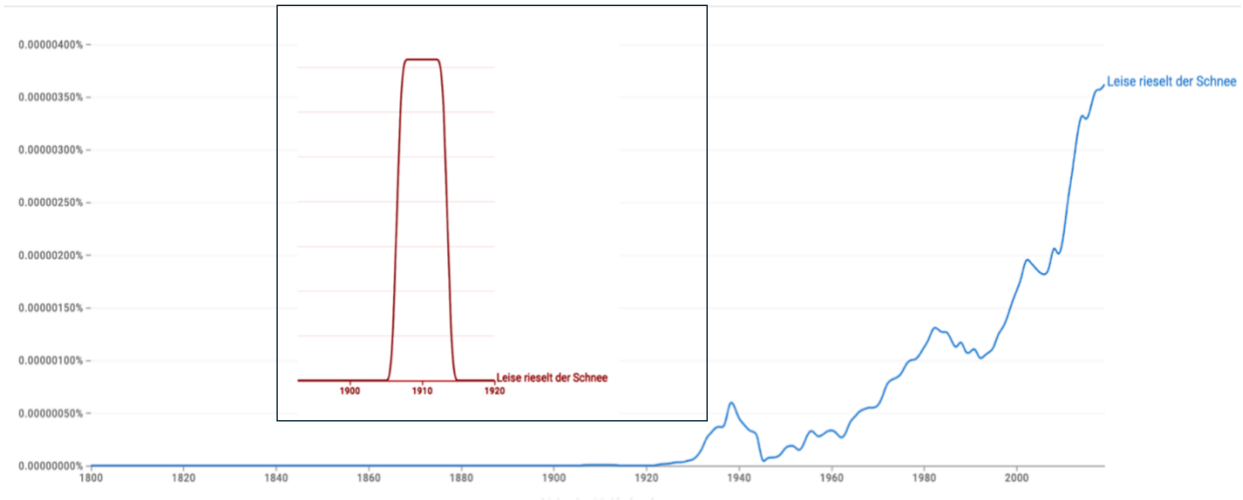

Figure 43. *Frequency curve of the multi-word sequence* Leise rieselt der Schnee *'Softly falls the snow' in the Google N-gram viewer between the years 1800–2000 (in blue) with an inset of the years 1900–1920 (in red)* [119]

Barring metaphorical uses, subjects that occur with *rieseln*, MEs or otherwise, are explored in the following section. The verb tends to show up in restricted collocations and both the *deTenTen13* and the *DWDS* corpora overlap in the set of subject collocates they return.

4.2.2.2 Subjects occurring with the verb rieseln *in the* deTenTen13 *and the* DWDS *corpora*

This subsection involves subjects that collocate with the verb *rieseln* based on the data provided by *SE*'s *Word Sketch* and the *DWDS Wortprofil*. Table 99 compares the most frequent subjects occurring with the verb *rieseln* 'flutter/trickle/drift' in *SE*'s *Word Sketch* and the *DWDS*

[118] Source: books.google.com/ngrams/info (last accessed March 7, 2024).

[119] The curve is, of course, not as pronounced when the years 1905–1916 are isolated.

Wortprofil, where possible precipitate types are bolded.[120] The subjects overlap greatly between these two corpora and both profiles return *Schnee* 'snow' as the most frequent subject of *rieseln*, (N= 2,560 and N= 773, respectively), with an incidence that is three to five times higher than the next most frequent subject, which is *Sand* 'sand' in both corpora. That being said, only in *Word Sketch* is the *logDice* score, which measures the collocational strength, also the highest for *Schnee* 'snow' (10.1). It is only the sixth highest in *Wortprofil* (8.1), where the highest *logDice* is provided by *Schneeflocke* 'snowflake' (9.3; after *Kalk* 'chalk/lime', *Flocke* 'flake', *Putz* 'plaster', and *leise* 'softly'). At the same time, *Schneeflocke* 'snowflake' is the third most frequent subject of *rieseln* in *Word Sketch* (N=203) with the fourth highest *logDice* (9.2). In comparison to the *Schnee* 'snow' precipitate, the *Regen* 'rain' precipitate is relatively infrequent with a weak-to-average *logDice* score (3.1 in *Wortprofil* and 5.5 in *Word Sketch*). Moreover, the *DWDS Wortprofil* returns *Regen* 'rain' as the 19th most frequent subject collocate.

Table 99. *LogDice scores and most frequent subjects collocating with verb* rieseln *in the* deTenTen13 *and in the* DWDS Wortprofil

Rank	***SE - deTenTen13***	Frequency	logDice	***DWDS Wortprofil***	Frequency	logDice
1.	***Schnee*** 'snow'	**2,560**	**10.1**	***Schnee*** 'snow'	**773**	**8.1**
2.	*Sand* 'sand'	491	9.4	*Sand* 'sand'	244	6.8
3.	***Schneeflocken*** 'snowflakes'	**203**	**9.2**	*Kalk* 'chalk/lime'	120	8.9
4.	*Staub* 'dust'	182	8.3	*Putz* 'plaster'	120	8.5
5.	*(leise)* 'softly'	169	9.6	*Staub* 'dust'	107	6.5
6.	*Flocke* 'flake'	157	9.0	***Wasser*** 'water'	**104**	**2.5**
7.	*Kalk* 'chalk/lime'	153	8.9	***Schneeflocke*** 'snowflake'	**83**	**9.3**
8.	*Putz* 'plaster'	144	8.5	*Pulver* 'powder'	81	7.6
9.	*Schauer* 'shudder/shower'	124	7.6	*Flocke* 'flake'	59	8.6
10.	***Regen*** 'rain'	**111**	**5.5**	**(19.) *Regen*** 'rain'	**31**	**3.1**

Based on the semantics, the subject *Schneeflocken* 'snowflakes' can be grouped with instances of *Schnee* 'snow', the only difference being that *Schnee* is a mass noun while *Schneeflocken* focuses on the unique participants that make up the mass. *Flocke* 'flake' primarily, but not exclusively, refers to snowflakes, which, in the *deTenTen13*, are furthermore modified with adjectives such as *weiß* 'white' (N= 75 out of 157 in the *deTenTen13*) and *erste* 'first' (N= 21) as well as a path (*vom Himmel* 'from the sky', N= 49 out of 157).

[120] The *DWDS Wortprofil* data for the verb *rieseln* were retrieved on February 29, 2024.

The fifth most frequent collocate in subject position is the manner adverb *leise* 'softly' (N=169), which is predominantly part of the fixed expression *leise rieselt der Schnee* 'softly falls the snow', as discussed above. *Leise* 'softly/quietly', as an adverb, cannot be an actual subject, but merely occurs in default subject position, i.e., at the beginning of the sentence. The actual subject, *der Schnee* 'the snow', has been extraposed in most of these instances (N=148), where *leise* occurs in subject position, the subject collocate is, indeed, *Schnee* 'snow'. The only subject that is predominantly metaphorical is *Schauer* 'shower/shiver', which, against all expectations, does not refer to showers of a certain precipitate, but to shivers running down one's spine. With the exception of *Wasser* 'water' in the *DWDS*, the top subject collocates for *rieseln* are almost identical between the two corpora.

Although the multi-word sequence and song title, *leise rieselt der Schnee* predominate, not all examples unambiguously quote or refer to the song title as does the example in (254).

(254) *… die Jungmusikanten versuchen mit "**Leise rieselt der Schnee**" ganz behutsam den Winter herbeizurufen.* (*DWDS – mv-michelwinnaden.de,* 'Music Club Concordia in Michelwinnaden')
… the young musicians are trying to gently incant winter with (the song) "**Softly the snow falls**".

The examples in (255) and (256) do not necessarily reference the song. In (256), if the speaker did indeed intend to use the song title, they could have kept "the original" *Leise rieselt der Schnee* followed by *an den Fenstern vorbei*. However, the word order is different, so that we do not know whether the speaker meant a play on words or not. The song title may simply have become conventionalized because of the popularity of the song. The corresponding frequent exposure to the song and, thus, the collocation may just as well have primed the speaker.

(255) ***Leise rieselt der Schnee*** *und verwandelt die Bergregion in ein zauberhaftes, weißglitzerndes Wintermärchen, das dazu einlädt, es zu erkunden.*
'**The snow falls softly** and transforms the mountain region into a magical white-glittering winter tale that invites (us) to explore it.' (*DWDS* – hotel-mohren.de)

(256) ***Schnee rieselt leise*** *an den weihnachtlich dekorierten Fenster vorbei, ich lausche Ella Fitzgerald und freue mich über den meinen Tannenbaum.* (*DWDS* – relleomein.de 'blog')
'**Snow is softly fallling** past the festively decorated windows, I'm listening to Ella Fitzgerald and take delight in my Christmas tree.'

Additionally, there are matches from the collection of the *DWDS* historical corpora (*Historische Korpora* [1465–1969]) which prove that this frequent, fixed expression cannot solely be attributed to the song and/or its lyrics, as they date back to before 1895, which is when the song was written. The year 1853 yields the oldest example returned by *DWDS* for the *Schnee &&*

rieseln search, but not for the *Schnee && rieseln && leise* search query. It is not clear from the 1853 concordance, however, that the snow drifted down from the skies, but rather from mountain tops. An explicit combination of the split [*Schnee rieseln*] construction, dated 1865, is given in (257), where the source is explicitly mentioned as *vom Himmel* 'from the sky'. This ties in with the findings from Chapter 3 that the predicate construction only infrequently occurs with a SOURCE path. However, since the utterance comes from the literary genre, it is possible that the author made a stylistic choice in employing the split type construction.

(257) *Ohne Aufhören* ***rieselte der Schnee*** ***vom Himmel****, sie wußten nur zu gut, daß er ihre Spur verwischte, - wie sollte man sie finden.* (*DWDS*: Ernst Keil (ed.): *Die Gartenlaube* 'The Garden Shed', Issue 13 (1865))
'**The snow was falling from the sky** without ceasing, they only knew too well that it covered their tracks – how to find them.'

In brief, the subject collocates of both corpora support that, nowadays, *rieseln* is predominantly used with snow. While a large number of [*Schnee rieseln*] collocates can probably be traced back to the song, the historical corpora and *Google N-gram* show that the use of the collocate (barring the manner adverb *leise* 'softly/quietly') dates back to before the song became popular. However, the returns in Table 99 also prove the existence of the [*Regen rieseln*] split type construction. Furthermore, despite the lower incidence of [*Regen rieseln*] collocations, Grimm's definition in (1) attests to *rieseln* collocating with other precipitate types where *Regentropfen* 'raindrops' and *Tautropfen* 'dewdrops' are mentioned first.

It is interesting that the *DWB*, in listing the precipitate types, specifies the individual units *Regentropfen*, *Schneeflocken*, and *Hagelkörner* 'rain*drops*, snow*flakes*, and hail*stones*' rather than the mass nouns *Regen* 'rain', *Schnee* 'snow', and *Hagel* 'hail'. In contrast, the other dictionaries do not necessarily specify the subject, but rather its consistency as *in vielen kleinen Teilchen* 'in many little particles'. Although, the *DWDS* notes *in feinen Tropfen niederfallen* 'fall down in fine drops' and the etymological entry for *rieseln* reads that, at least in Middle High German, the meaning of *riselen* appeared to specify the rain precipitate rather than snow (*tröpfeln*, *leise regnen* 'drip, rain softly', cf. [c] above). The *DWDS* further details that *nieseln* is possibly an *anlaut* variant of *rieseln* based on *nass* 'wet' or *nässeln* (archaic) 'to get slightly wet', which original meant 'falling inaudibly (of dying leaves)' and later 'of fine rain'. It is also listed in the *DWDS* dictionary as *leise regnen* 'rain softly' and *ein wenig nässen* 'wet/moisten a little'. [121]

[121] Source: www.dwds.de/wb/nässeln (last accessed, April 01, 2024).

Additionally, the *DWDS* mentions *Nieselregen* 'drizzle' meaning *feiner Sprühregen* 'fine spray rain' (20th century), in comparison with the older, declining *Rieselregen* 'trickle rain', which was more frequently used in the 18th century.[122] Accordingly, let us explore the *Regen rieseln* collocation. Table 100 gives the frequencies for the search queries for the precipitates with *rieseln*] in the *deTenTen13* for both the PREC1 and the PREC2 word order as well as a search in three *DWDS* metacorpora. The exact CQL for the *deTenTen13* search was "[word="Regen"][lemma="rieseln"] within <s/>" where the precipitate was either *Regen* 'rain' or *Schnee* 'snow' and where the word order was inverted to obtain PREC2 results. The search query for *DWDS* was "*Regen rieseln*" and "*Schnee rieseln*", here, too, for the inverted word order as well. Take note that these collocates differ from the subject collocates in Table 99, which are returns provided by the *Word Sketch* function and *Wortprofil*. The search queries in Table 100 did not allow for any modifying items between the precipitate and the verb. As already stated, [*Schnee rieseln*] (N= 434) is roughly eight times as frequent as [*Regen rieseln*] (N= 55) within the *deTenTen13*. The total of the *DWDS* results show that *Schnee* is roughly five times as frequent as *Regen* (*Regen* N= 95 vs. *Schnee* N= 484) with *rieseln*, however, the numbers differ depending on which metacorpus provided the results. For the historical corpora, the [*Regen rieseln*] construction has a higher incidence (N= 57) than [*Schnee rieseln*] (N= 11) with the incidences of 5:1 inverted. Only the more recent metacorpora, namely *WebXL* and the *Gegenwartskorpora* 'contemporary corpora' have much higher incidences of [*Schnee rieseln*]. This is another indicator that attests to the change in collocational preference for *rieseln*.

[122] Source: www.dwds.de/wb/nieseln (last accessed, April 01, 2024).

Table 100. *Frequencies of the* Regen/Schnee rieseln *collocations in the* deTenTen13 and the DWDS corpora *for PREC2 and PREC1 word order*

Corpus (and subcorpora)	Precipitate	PREC2	PREC1	Total
deTenTen13	*Regen* 'rain'	**52**	**3**	**55**
	Schnee 'snow'	**399**	**35**	**434**
DWDS				
Historical Corpora	*Regen* 'rain'	54	3	57
WebXL	*Regen* 'rain'	29	1	30
Contemporary Corpora (collection)[123]	*Regen* 'rain'	8	0	8
Total *Regen*		**91**	**4**	**95**
Historical Corpora	*Schnee* 'snow'	10	1	11
WebXL	*Schnee* 'snow'	410	34	444
Contemporary Corpora (collection)	*Schnee* 'snow'	24	5	29
Total *Schnee*		**444**	**40**	**484**

The oldest return for the lemma *rieseln* in the *DWDS* historical corpus that refers to light rain dates back to 1629, as given in (258). Here, the "stronger" ME predicate *regnen* 'rain' is juxtaposed with the less forceful predicate *rieseln* 'drizzle' (*nicht geregnet .. sondern geriselt* 'not rained but drizzled/trickled'), but a noun collocate *Regen* 'rain' is not provided.

(258) *Matthesius, meldet an einem Ort … **habe es** dieselbe Zeit am Tage **nicht geregnet/sondern nur** des Nachts **geriselt**/ damit selbige Arbeit nicht verhindert würde*.
(*DWDS*: Paul Kentz: *Güldener Handwercksboden*. 'Golden trade' Leipzig, 1629)
'Matthesius, reports in one place … (reportedly) **it did not rain** at the same time during the day/ **but only drizzled** at night/ so that the same work would not be hindered.

Where *Regen rieseln* is used in more contemporary utterances, the sense of "fine" rain is retained and many of the concordance lines returned from the various corpora appear to collocate with the adjectives *fein* 'fine' and *leicht* 'light'. This is also true for (259) and (260), which stem from a 1916 novel and a 2013 magazine, respectively.

(259) ***Ein feiner Regen rieselte** wie Nebel **nieder**.* (gutenberg.de [1916])
'**A fine rain fell down** like fog.'

(260) *Auch wenn ich jetzt zum Fenster hinausschaue, **rieselt ein leichter Regen vom Himmel**.*
(www.cherry.ch [2013])
'Even when I look out the window now, **a light rain is falling from the sky**.'

Both of these examples contain a PATH (*nieder* 'down' and *vom Himmel* 'from the sky'), but for many other examples, it appears that the sound component of the verb's semantics is more

[123] A list of the *DWDS* metacorpus *Gegenwartskorpora mit freiem Zugang* 'Contemporary corpora with free access' can be found here: https://www.dwds.de/d/korpora/dwdsxl (last accessed April 1, 2024).

prominent, as in (261) and (262). This is emphasized by the modifying material, i.e., the alliteration of *rauschen* 'rush' and *rieseln* in (261) and the adverb *leise* 'softly/gently/quietly' in (262). Example (261) sounds rather poetic, but it is "trivial" weather talk written on the message board of a TV series. It appears that where the sound component is presumably more salient, the path is not as relevant. This observation may provide a further avenue of research: to what extent are modifiers profiled and do they cancel each other out or is there some sort of hierarchy. Furthermore, example (262) contains the only GOAL path in all of the returned concordance lines.

(261) *Draußen* ***rauschte und rieselte*** *gestern so herrlich versunken* ***der Regen***… (tatort-forum.de)[124]
'Outside **the rain rustled and trickled** so beautifully yesterday.'

(262) ***Der Regen rieselt*** *gerade* ***leise auf die Balkonkästen****, eine gute Gelegenheit, das Bett aufzusuchen.* (*DWDS*: *Sigitaria's Blog*)
'**The rain is** currently **falling quietly onto the balcony boxes**, a good opportunity to go to bed.'

The utterance in (263) presents an example of a quasi-transitive, "tripartite" collocation in a similar manner to the [*es geben* precipitate] construction. However, here the pronoun *es* 'it' precedes a verb (*rieseln*) which agrees in number with the postposed plural subject *laue Regen* 'mild rains', while in the [*es*+*gibt* precipitate] construction the verb *gibt* 'gives' agrees with singular pronoun *es* 'it'. Even so, there are two arguments, one light and one heavy and the predicate, as opposed to the predicate in the [*es geben* precipitate] construction, is also heavy. The function of *es* 'it' here is an abstract scene-setting one within which the "action" or process occurs (cf. Smith 2002). Using a "camera metaphor", Smith explains this as the human tendency to "take a wide-angle view on a scene before focusing (…) upon specific entities or actions within that space" (2002: 79). Additionally, as a text from the poetry genre, the utterance in (263) is highly marked, the preceding line ends in *Segen* 'blessing', which obviously rhymes with *Regen* 'rain'.[125] To achieve that rhyme, the author employed a scene-setting *es* 'it' to fill the clause initial position, which usually translates to 'there' in English evoking a spatial feel. There is no structural parallelism otherwise between the two clauses in (263). Furthermore, the author uses other similar scene-setting constructions with clause-initial *es* 'it' in the poem (*es schweigt das aufgewühlte Meer* 'the rolling sea is silent', lit. 'it silences the rolling sea', *es scheint der stille*

[124] *Tatort* 'crime scene' is a popular German TV show.

[125] The complete poem can be found on www.lyrikkompass.de/poems/21089 (last accessed on September 21, 2024).

Mond 'it shines the quiet moon' where, however, the respective verb is inflected for third person singular) in addition to more "conventional" *es* 'it' predicate constructions like *es blitzt* 'it flashes', i.e. 'lightning flashes/strikes'. Like (261), example (263) also includes the ideophone *rauschen* 'rush/rustle', but in combination with clouds.

(263) *Der Wolken Fülle rauscht; **es rieseln** laue **Regen**.*
'The clouds' fullness rushes; **it trickle** mild **rains**.'
(*DWDS*: Ludwig Gotthard Kosegarten: *Poesieen* 'Poesies'. Vol. 2. Leipzig, 1798)

It is not possible to say *es rieselt* alone, without any further context, and for the listener to immediately infer that the speaker is referring to snow. Still, the strong collocation with *Schnee* 'snow' indicates a high probability that, provided the right context, a speaker could identify the subject matter as the *Schnee* 'snow' precipitate. It is, however, possible and commonplace to say *es nieselt* 'it drizzles' and the recipient/interlocutor will identify the precipitate as rain and not snow. There are only four examples of *nieseln* 'drizzle' where the predicate collocates with *Schnee* 'snow' in all of the *DWDS* corpora combined and in these examples, as given in (264), the object *Schnee* 'snow' is obviously always specified. It is possible that where *nieseln* 'drizzle', rather than *rieseln*, collocates with *Schnee* 'snow', the precipitate itself is "wetter", closer to *Schneeregen* 'sleet' than to dry *Pulverschnee* 'powdery snow'.

(264) *Eisiger **Schnee nieselt** auf die Neugierigen in Maske und Zivil (…)*
(*DWDS – Die Zeit* 'The Times newspaper')
'Icy **snow is drizzling** onto the curious (people) in masks and plainclothes.'

At the same time, there are just as many examples that juxtapose *Schnee* and *nieseln,* as exemplified in (265) and (266). Take note that (265) gives both argument-predicate collocates (*Schnee rieseln* vs. *Regen nieseln*) while (266) contains a negative assertion regarding the *Schnee* 'snow' precipitate, in the first clause, followed by a "replacement ME" in the second clause: the predicate *nieseln* 'drizzle' needs not specify the *Regen* 'rain' precipitate.

(265) *Auch wenn es nicht **der Schnee, der gerieselt**, sondern **der Regen** war, **der** leise **genieselt** hat, so war die Stimmung doch fröhlich-festlich – jauchzet! Frohlocket!*
'Even if it was not **the snow that was fluttering**, but **the rain that was drizzling** softly, the atmosphere was cheerful and festive – shout for joy! Rejoice!'
(*DWDS Webkorpus*: *Im Stadion* 'at the stadium' – boogiebaron.de, 2020)

(266) ***Schnee** ist hier immer noch nicht in Sicht, stattdessen **nieselt es**…*
'Snow still isn't in sight here, instead **it is drizzling…'**
(*DWDS Webkorpus*: *Arianas Adventslesen* 'Ariana's advent reading', 2017)

Last but not least, in Chapter 3, I pointed out the fact that German can easily forego any kind of additional (e.g., adverbial) modification, such as manner or quantity, by applying its well-known proclivity for nominal compounding (closed compounds, no hyphen or spaces used). Weather events pose no exception and, accordingly, there exists the compound *Nieselregen* 'drizzle rain'. Take note that one also finds *Rieselregen* and *Rieselschnee* entries in the *DWDS*, but not in the *Duden*.[126] It appears that *Rieselschnee* is often used for artificial snow. The *DWDS* corpora return 15 instances of *Rieselschnee* where seven refer to a children's theatre company named *Rieselschnee und Klingelglöckchen* and their identically named music CD. The *DWDS* dictionary defines *Rieselregen* as 'fine rain' or 'spray rain'; *Rieselschnee* refers back to the *Rieselregen* entry without being afforded its own definition, but, in return, there is no such thing as *Nieselschnee* 'drizzle snow'.[127] Nevertheless, a *Google* search provides roughly 880 instances of *Nieselschnee*, two of which are given in (267) and (268). Again, the word choice indicates that the snow is wetter than "usual".

(267) *Auch **bei** 6 Grad und **Nieselregen** (zwischendrin mehr **Nieselschnee**) kann man Outdoor-Möbel draußen attraktiv fotografieren.* (heldenstreich.com – an agency for employer branding)
'Even **at** 6 degrees and **drizzle rain** (in between more **drizzle snow**) you can take attractive photos of outdoor furniture outside.'

(268) *Trotz eisiger Temperaturen und **Nieselschnee** ging es im Zwiebellook und Laufschuhen Richtung Volkspark Schöneberg.* (mitvergnuegen.com 'with pleasure' – a digital city magazine)
'Despite icy temperatures and **drizzle snow** we headed towards Volkspark Schöneberg sporting the onion look and running shoes.'

Nieselregen is used much more frequently than *Niesel* alone, which is neither part of the *Duden* nor of the *DWDS*, but, nevertheless, returns numerous concordance lines in the *DWDS* (N= 7,096 in total in 22 corpora). The *Niesel* returns probably contain a fair amount of false returns, where *Niesel* is the last name of a person. For this reason, I left out the counts of a few corpora, such as the *Jurakorpus* 'law corpus' (N=1,929), whose returns I scanned and which appear to contain the proper name *Niesel* only rather than actual weather events. In the *DWDS* dictionary, the definition of *Nieselregen* is given as *nieselnd niedergehender Regen* 'rain that goes down drizzling'. The *deTenTen13* returns 2,455 concordance lines for a simple *Niesel* (i.e., the exact

[126] For both entries, the *DWDS* references the *eWDG*, the electronic version of *Das Wörterbuch der deutschen Gegenwartssprache* 'The dictionary of the German contemporary language', which was compiled between 1952 and 1977 at *The German Academy of* Sciences in Berlin.
[127] www.dwds.de/wb/Rieselregen (last accessed February 15, 2024).

word) search query. These returns contain many utterances where *Regen* 'rain' is distinguished from *Niesel* 'drizzle', as in the weather report in (269).

(269) *In der Nacht zu Donnerstag **fällt** in Süddeutschland gebietsweise **Regen** **oder** **Niesel**, im Bergland fällt etwas Schnee.* (*Die Zeit* 'The Times online newspaper')
'On Thursday night, **rain** **or** **drizzle** **falls** regionally in parts of southern Germany, and some snow falls in the mountain regions.'

The *DWDS*, too, provides corpus examples where *Niesel* is used as a noun in opposition to *Regen* 'rain', as in (270).

(270) *In der Nacht **fällt** bei 7 bis 4 Grad etwas **Regen** **oder** **Niesel**.* (*DWDS*: *Rhein-Zeitung*, 07.03.2013 'Rhenish newspaper')
'At night **falls** some **rain** **or** **drizzle** at 7 to 4 degrees.'

Obviously, both sources are newspapers and, thus, adhere to some level of formality so that the use of *Niesel*, instead of *Nieselregen*, cannot simply be ascribed to informal language use. The *Deutscher Wetter Dienst* 'German Weather Service' (or *DWD* for short) itself does not (yet) list *Niesel*, but only *Nieselregen* and the verb *nieseln*, which both link to *Sprühregen* 'spray rain' instead of providing their "own" definition.

In keeping with rain-compounds, there are many different types of rain in addition to *Nieselregen*, as defined by the *DWD* and listed under *Niederschlag* 'precipitation'.[128] These are, for example, *Sprühregen* 'spray rain'*, Platzregen* 'downpour', *Landregen* lit. 'land rain' or *Dauerregen*, with both the latter meaning 'constant rain', *Starkregen* 'heavy rain' (which in Germany only happens in the mountains or during thunderstorms), *Eisregen* 'freezing rain'*, Monsunregen* 'monsoon rain'*, Tropenregen* 'tropical rain'*,* and *Blutregen* 'blood rain', which is rain that appears in the middle of Europe and has been contaminated by dust from the Sahara. In these compounds, the first noun typically describes the type or MANNER of the rain coming down. Since both the argument itself encodes the ME and simultaneously modifies it, these compounds are certain to appear in an argument construction. These rain compounds are listed in the *DWD*, but they are not just part of the meteorological professional terminology, they are also part of children's vocabulary, or at least children are being introduced to them, as shown by a clip of *Piggeldy und Frederick*. *Piggeldy und Frederick* is a popular children's cartoon, recommended for children ages three and up, which originally aired from 1973–1990, but has

[128] www.dwd.de/DE/service/lexikon/Functions/glossar.html?lv2=101812&lv3=101914 (last accessed February 15, 2024).

frequent reruns on TV. In the show, the protagonists, two pig brothers, make sense of the world. The older brother, Frederick, tries to answer his younger brother, Piggeldy's, incessant questions about any conceivable subject, e.g., concepts such as *love* and *patience*, but also more tangible things, such as the weather. In *Der Regen* 'The Rain', Piggeldy asks Frederick *Was ist Regen* 'what is rain', which Frederick answers by listing all types of rain.[129] The relevant bit of their conversation is given in (271).

(271) Frederick: *Es gibt **Platzregen**, **Dauerregen**, **Sprühregen**, **Nieselregen**, **Eisregen**, **Landregen**, **Gewitterregen**, **Sommerregen**, **Winterregen** und **Bindfadenregen**.*

(272) Frederick: 'There are **cloudbursts**, **constant rain**, **spray rain**, **drizzle rain**, **freezing rain**, **steady rain**, **thunder showers**, **summer rain**, **winter rain**, and **pouring rain** [lit. 'string/thread rain']."

Obviously, (271) contains an [*es gibt*] construction; in this case, however, the construction involves an existential usage, since Frederick enumerates different types of rain that exist in general. He does not make any predictions about the ongoing weather situation as the two go on their customary little walk. Nonetheless, it suddenly starts raining, which the narrator remarks on, as in (273).

(273) *Narrator: Und kaum hatte Piggeldy dies ausgesprochen, **fielen dicke, runde Tropfen auf die beiden hinab**.*
Narrator**:** And as soon as Piggeldy had said this, **thick, round drops fell down onto the two of them**.'

Take note how the narrator does not resort to a predicate construction (e.g., *as soon as Piggeldy had said that it started raining onto the two of them*), but he uses an argument construction with *fallen* 'fall'. It is possible that the writers chose this construction to "better" suit the argument *dicke, rund Tropfen* 'thick, round drops' instead of *Regen* 'rain'. However, there are no syntactic or semantic constraints that would disallow the predicate construction in combination with the given argument (*es fing an dicke Tropfen auf die beiden hinabzuregnen* 'it started raining thick, round drops onto the two of them').

Aside from being a children's show, a variety of audio and video clips of *Piggeldy und Frederick*, including *Der Regen*, has also been integrated in the immersive language learning

[129] The video clip can be found on YouTube (www.youtube.com/watch?v=vvwYQjygGPc, last accessed April 06, 2021)

website *Yabla*, where many of the clips concerning the weather are listed as "beginner".[130] Frederick's explanation contains most of the types of rain common in Germany and also adds a few extra "types", i.e., the last three, which are more frequent in common language and not so much a part of meteorological jargon. Furthermore, *Sommerregen* 'summer rain' has its own entry in the *Duden* as opposed to *Winterregen* 'winter rain'. Both compounds have an entry in the *DWDS* dictionary, where the entry for *Sommerregen* 'summer rain' references the *WDG* (cf. fn.126) as updated by the *ZDL* (*Das Zentrum für digitale Lexikographie der deutschen Sprache* 'The Center for the digital lexicography of the German language') in 2023. The *Winterregen* 'winter rain' entry appears to be more recent, since it only references a *ZDL* article dated 2023, but does not make reference to the older *WDG*. Lastly, Frederick mentions *Bindfadenregen* 'thread/string rain', which *Duden* defines as *anhaltender, strömender Regen* 'lasting, pouring rain' and which is probably derived from the expression *Es regnet Bindfäden* 'it rains threads/strings'.[131] This expression is briefly discussed again in Section 4.5, which involves figurative MEs.

In summary, the [precipitate *rieseln*] construction is a split construction that encodes the ME in the argument. Nowadays, it does so primarily, but not exclusively, with the *Schnee* 'snow' precipitate, while historical documents indicate a past preference for the *Regen* 'rain' precipitate (and possibly other precipitate types like *Hagel* 'hail'). It is possible that the popularity of the song *Leise rieselt der Schnee* 'Softly falls the snow' has led to a predominance of the *Schnee rieseln* collocation. At the same time, a development of the *anlaut* variant *nieseln* has come to collocate more frequently with the *Regen* 'rain' precipitate. Where *Regen* 'rain' collocates with *rieseln* in more contemporary utterances, the sound component of the definition may prevail. Both *riesel-* and *niesel-* have also shown to compound with *Regen* 'rain' and *Schnee* 'snow', but the latter represents informal use with no entry in either of the consulted dictionaries. Both the use of the split construction and the compounds appear to rely at least partly on the quality of the precipitate, i.e., where the precipitate is dry, there is a preference for *riesel*; where the precipitate is more wet, there is a preference for *niesel*. Additionally, the component of sound comes into play where *Regen* 'rain' is concerned. The examination of the compounds *Rieselregen* and *Nieselregen* led to a brief exploration of other compounds, which

[130] https://deutsch.yabla.com/Piggeldy-und-Frederick-series-1108 (last accessed April 06, 2024)

[131] Source: www.duden.de/rechtschreibung/Bindfadenregen (last accessed April 01, 2024). The French have the very similar expression *il pleut des cordes*.

occur in professional meteorological jargon as well as in children's language and vocabulary directed at language learners. While these compounds themselves will clearly occur in an argument construction, with [*es geben*] or [*fallen*], they simultaneously encode a facet of the weather event itself, be it manner (*Sprühregen* 'spray rain'), temporality (*Dauerregen* 'continuous rain'), seasons (*Sommerregen* 'summer rain'), or concomitant MEs (*Gewitterregen* 'thunderstorm rain'). None of these concomitant events are linguistically crucial or obligatory, but speakers nevertheless customarily make mention of these modifications, as these facets may be crucial in other regards. Thus, while the [precipitate *rieseln*] constructions are unmarked regarding their grammatical make up—they are simple argument+predicate constructions with semantically heavy+heavy components—they are marked with regard to their usage in view of everyday weather events. The next section involves similarly syntactically unmarked constructions, in the shape of transitive clauses, where causal agents are employed as entities responsible for specific MEs. These may be generic, such as 'the world' or 'the sky', but the focus here is on two agents specific to the German-speaking cultural areas, i.e., *Frau Holle* 'Mother Hulda' and *Petrus* 'Saint Peter'.

4.3 Other Marked Constructions

4.3.1 Transitives with causal agents

This section is concerned with agentive ME constructions, which represent only a minor ME schema in German. Nevertheless, it is a common strategy in some languages to employ a natural agent as the source or causer of the ME, such as the world or a place, as in (273), which Givón (2001) categorizes as "dummy subjects".

(273) *Id-dunya ti-shti* [Palestinian Arabic, Givón 2001: 119]
the-world 3SG.FEM-rain
'it is raining' (lit.: 'the world is raining')

Eriksen et al. (2010: 577) give a number of examples of this construction type and refer to them as "locational subjects", "temporal subjects", such as 'day', 'time', and "atmospherical subjects", such as 'sky' or 'weather'. They point out that examples like these, especially where the meteorological subject is 'weather', are commonplace. Such examples can indeed easily be

found in English as well as in German: *Das Wetter ist kalt/regnerisch/stürmisch* 'The weather is cold/rainy/stormy'. Take note, though, that *das Wetter regnet* 'the weather is raining' is odd at best, although examples can, of course, be found.[132] Other atmospherical agents that can be found for precipitation events in German are the clouds or the sky.

The lack of clearly distinguishable or distinct participants in MEs seems to be the underlying cause for the diversity and the unpredictability of the syntactic forms MEs take. One strategy, which languages employ to account for this perceived lack is the introduction of agentive subjects, especially deities, as in the Latin example in (274).

(274) *Juppiter pluit*. [Latin, Bauer 2000: 103]
Jupiter rain.3SG.pres
'Jupiter rains.' / 'Jupiter sends rain.'

Other attested agentives from Christian times in Latin were, for example 'the sky' (*caelum*) or 'the Lord' (*Dominus*) (Bauer 2000: 104). As already established, German allows for the evocation of generic *Wettergötter* 'weather gods', as in (275), where the coordinators of some unnamed event express their hopes that the weather god will be favorably inclined towards them, which, in all likelihood, means that they hope that it will not rain. However, the particular ME itself is not spelled out.

(275) *Natürlich hoffen wir, dass* ***uns*** *wie in den vergangenen Jahren* ***der Wettergott gewogen bleibt*** *und seinen Teil zum Verweilen der Besucher beiträgt.* (slowfood-sachsen.de 'slowfood Saxony')
'Of course we hope that, as in previous years, **the weather god will remain kind to us** and do his part to ensure that visitors will spend some time.'

In example (276), in contrast to (274), where Jupiter "himself" rains, the ME, *Regen* 'rain', is mentioned and describes the weather condition that the speaker's dog is not willing to abide; the weather god here *lässt es regnen* 'lets it rain', where *lassen* 'let' implies causality.

[132] ***Wenn das Wetter regnet,*** *is* [sic] *das Holz halt nass*. 'The wood is wet then **if the weather is raining.**' (Source: twitter.com/saucychilisauce/status/1528125520675803140 last accessed April 19, 2024) or ***Wenn das Wetter regnet***, ist es von *größter Bedeutung zu wissen, dass Ihre Ausrüstung trocken gehalten wird*. 'It is crucial to know that your gear is going to be kept dry **when the weather is raining**.' (Source: alltricks.de, description of a backpack) (last accessed April 19, 2024).

(276) *Ist **der Wettergott** erzürnt und **lässt es tagelang regnen**, dann setzt unser Luxushund keine Pfote vor die Tür, was er gar nicht abkann ist Wasser von oben.* (stoanahof.de – dog breeder)
'If **the weather god** is angry and **lets it rain for days**, then our luxury dog won't put a paw in front of the door, what he cannot stand at all is water from above.'

Deities are obviously not an intrinsic part of an ME, but their inclusion may result from a linguistic reflex to re-frame a naturally occurring environmental situation as the result of a causal chain, as in a transitive clause. Be that as it may, if we consider actual language use, we must include not just "trivial" examples of the argument type kind, but also idiosyncratic and expressions that vary by culture. Two of the canonical German cases that come to mind are *Frau Holle* 'Mother Hulda' and *Petrus* 'Saint Peter'. *Frau Holle* is a pre-Christian legend and deity in her own right (not just in Germany) and is thought to appear mostly during the winter months, more specifically, the *Raunächte* (the twelve nights after Christmas). She is associated with many things, but in Germany she is probably mostly recognized from the fairy tale of *Frau Holle*, as collected by the Brothers Grimm where she is responsible for making it snow on earth and has, thus, come to be associated with snowfall. On the other hand, *Petrus*, who is believed to be responsible for the weather in general, is more closely associated with rainfall – or with dry and sunny weather when he is agreeable. The following subsections explore MEs with these two mythical weather agents, starting with *Frau Holle*.

4.3.1.1 Frau Holle *'Mother Hulda'*

The tale of *Frau Holle* 'Mother Hulda' is a German fairy tale collected by Jacob and Wilhelm Grimm in their *Kinder- und Hausmärchen* 'Children's and Household Tales' (1812).[133] However, *Frau Holle* is an important legendary creature from Germany in her own right and the earliest written mention of *Frau Holle* is found in the decrees of Bishop Burchard von Worms, written between 1008 and 1012 (Kollmann 2012). In traditional folklore, *Frau Holle* is responsible for making it snow by making her bed; that is, she shakes her feather bed and the feathers turn into snow as they fall to earth.[134] The saying that *Frau Holle schüttelt ihre Betten aus* 'Mother Hulda

[133] *Frau Holle* (ATU 480) is the 24th story in the first volume of the book, published in 1812, as part of *Children's and Household Tales*.

[134] The story originates from Central Germany. In southern Germany and the Alps, she is known chiefly as *Perchta*. Most people familiar with the fairy tale are probably not aware that in Germanic Pre-Christian folklore, as presumed by Jacob Grimm, *Hulda*, *Holda*, *Holle* and *Holla* all denoted a single goddess, who rules the weather, but is also responsible for an array of other activities. She is generally perceived to be benevolent (the adjective *hold* is archaic for 'lovely/fair').

is shaking out her beds' has come to mean that it is snowing without necessarily having to mention *Schnee* 'snow' itself.

A simple search for *Frau Holle* returns 13,424 instances in the *deTenTen13* and 13,587 in the *DWDS* corpora. This does not mean, of course, that these returns actually encode or even imply an ME. The frequencies for the search query with the words *Frau Holle* and *Schnee* 'snow' and the verb lemma *schneien* 'to snow' are given in Table 101. The search query allows for up to eight items between the target words. The incidences do not differ much between the word orders, but the argument precipitate *Schnee* 'snow' appears much more frequently than the predicate precipitate *schneien* 'snow'.

Table 101. *Frequencies of* deTenTen13 *search queries containing* Frau Holle *and the* Schnee *'snow' precipitate*

CQL	Frequency
[word="Frau"][word="Holle"][]{0,8}[word="Schnee"]	285
[word="Schnee"][]{0,8}[word="Frau"][word="Holle"]	287
[word="Frau"][word="Holle"][]{0,8}[lemma="schneien"]	169
[lemma="schneien"][]{0,8}[word="Frau"][word="Holle"]	138
Frau Holle && *lassen* 'let' & *schneien* 'to snow' (DWDS)	134

In comparison, a search for *Frau Holle* && *lassen* 'let' & *schneien* 'to snow' in the *DWDS* metacorpora returns 134 concordance lines in total, most of which are in the *WebXL* corpus. The examples in (277) and (278) were selected from the returned results where the former contains a collocation with the argument *Schnee* 'snow' and the latter a collocation with the predicate *schneien* 'snow'. Furthermore, as in example (276), both contain the causative verb *lassen* 'let'. While (277) appears more to be like informal weather talk, as indicated by the informal *fette Ladung Schnee* 'fat load of snow' and the use of *ne* instead of *eine* 'a' (other than being narrated in first person plural *wir* 'we'), (278) originates from a newspaper article. Additionally, (277) contains the canonical, light predicate *fallen* 'fall'.

(277) *Nachdem wir am Mittwoch unseren ersten Ski-Spass dieser Saison bei Regen in Sankt Andreasberg hatten,* ***hat*** *nun* ***Frau Holle ne fette Ladung Schnee*** *über dem Harz* ***fallen lassen***. (tachales.de)
'After we had our first skiing fun of the season at rain in Sankt Andreasberg on Wednesday, **Mother Hulda has** now **let a fat load of snow fall** over the Harz Mountains.'

(278) ***Frau Holle ließ es*** *am Samstag in Dormagen* ***schneien***. (ngz-online.de 'Neuss-Grevenbroich Newspaper')
'**Mother Hulda let it snow** in Dormagen on Saturday.'

The above examples directly associate *Frau Holle* with the *Schnee* 'snow' precipitate; however, the ME can also be expressed by referencing what happens in the fairy tale, as in (279).

(279) ***Frau Holle*** *war seeehr fleißig und* ***hat*** *ordentlich* ***ihre Betten*** *über uns* ***ausgeschüttelt***. (sparen-tierisch-gut.blogspot.ch 'saving-beastly-good')
'**Mother Hulda** was veeery industrious and **has** properly **shaken her beds** on (lit. over) us.'

This example is "typical" in that the blog entry, from a dog's perspective, contains the familiar action: *Frau Holle* shaking her feather beds in order to make it snow. The blog entry is titled *Spaß im Schnee* 'fun in the snow' and it is accompanied by pictures of two dogs in the snow so that the ME is contextualized. The double reduplication of the letter "e" in *seeehr* 'very' and the additional use of the adverb *ordentlich* 'properly/thoroughly' imply that it snowed heavily without having to use a canonical ME or other modifier. Similarly, instead of having to resort to the metaphorical "shaking out her beds", it is possible for speakers to solely indicate that *Frau Holle* was *fleißig* 'industrious' or 'busy', as in (280).

(280) *In der letzten Nacht* ***war Frau Holle fleißig****. Mit gut* ***20 Zentimeter Neuschnee*** *hat sie das Allgäu in eine Märchenlandschaft verwandelt*. (kolping-familienurlaub.de 'Kolping family vacation')
'**Mother Hulda was industrious** last night. With a good **20 centimeters of fresh snow**, she has transformed the Allgäu (author's note: region in Bavaria, Germany) into a fairytale landscape.'

In (280), the subsequent context specifies the amount of fresh snow that has fallen. Nevertheless, an explicit mention of *Schnee* 'snow' is not mandatory. As evidence, the sentences in the blog entry, which precede and follow the use of *Frau Holle* in (281), do not contain any mention of the argument *Schnee* 'snow' or the predicate *schneien* '*snow*(ing)'. However, to the majority of German speakers, it will be clear that the planned hike did not take place due to an unexpected snowfall.

(281) ***Die Wanderung wurde*** *aber* ***dank Frau Holle in einen Kletterkurs*** *im Klettertreff in Thun* ***umgewandelt****.* (skiklub-zweisimmen.ch 'Ski Club Zweisimmen')
'**Thanks to Mother Hulda**, **the hike was turned into a climbing course** at the climbing club in Thun.'

Another fixed expression, other than *Frau Holle ist fleißig* 'Mother Hulda is being industrious', is *Frau Holle hat es gut gemeint* 'Mother Hulda meant well', as in (282):

(282) *Auch* ***Frau Holle meinte es gut*** *mit unserem Termin zum Winterwandern.* (oevp-vorchdorf.at)
'**Mother Hulda**, too, **meant well** with our date for a winter hike.'

Here, too, there is no mention of the actual ME. The subsequent paragraph of the blog entry relates how the hikers experienced an 'idyllic winter landscape'. From the sentence containing only *Frau Holle* meaning well, the reader can infer that snow has fallen. The ensuing description of the landscape provides enough contextualization to suggest that it snowed, furthermore, a decent amount.

The data have revealed that *es gut meinen* 'to mean well' is a figure of speech that collocates frequently with weather agents. That being said, there is a difference in semantic prosody depending on the agent. Where *Frau Holle*—the agent responsible for snowfall—is concerned, 'meaning well' usually implies a lot of precipitation. Where *Petrus* is concerned, it usually means that there was no precipitation and that the day was, at the very least, dry, if not sunny, as shown in the next subsection.

4.3.1.2 Der heilige Petrus *'Patron Saint Peter'*

Wetterheiliger 'weather saint' is what in folklore people call a Christian saint that is invoked to ask for favourable weather or to ward off famine. For example, *Saint Anne* is a protector from storms; the brothers *John* and *Paul* are protectors from hail (so called *Hagelprozessionen* 'hail processions' took place in their honour, which are nowadays called *Eucharist Processions*), thunderstorms, lightning, and rain; and, last but not least, there is *Petrus*, who stands as a cultural-historical successor to the Roman weather god, *Janus*. *Petrus* is often depicted holding a key. This goes back to the Gospel of Matthew 16:19, when Jesus says to Peter, "I will give you the keys of the kingdom of heaven, and whatever you bind on Earth shall be bound in heaven, and whatever you loose on Earth shall be loosed in heaven." Traditionally, Peter is thus understood to be the gatekeeper of heaven.[135]

[135] According to the theologian Theodor Lohmann (1960), the connection between Janus and Peter might be as simple as the weather usually being "a matter for the boss" in the pantheon. Naturally, Peter, as the most prominent apostle and the first leader of the early church, is thus perceived to be akin to Indra, Zeus, or Thor. (The principal deity of the Germanic pantheon is Odin, though, and not Thor.) Other parallels that Lohmann discusses are the "rock motif" (Matthew 16:18) where the Lord speaks to Peter "I say unto thee, that thou art Peter, and upon this rock will I build my Church; and the gates of hell shall not prevail against it." Weather gods are often portrayed as sitting on mountains high up in the skies where rain and thunderstorms come from. Paradoxically, to get to *Frau Holle*, who is sitting in the clouds and letting it snow down, the protagonist in the fairy tale jumps deep down into a well. Lohmann discusses a few other parallels but although his explanations highlight parallels between Janus and Peter, he does not present any actual proof for his hypotheses why it was Peter who became responsible for the weather. The connection remains purely speculative.

In folk belief, especially in German-speaking regions, *Petrus* is responsible for the weather and more specifically for rain. It is not clear why *Petrus* has been put in charge of weather, but possibly it is, because he is up in heaven and regulates the opening and closing of the "Pearly Gates". When he opens the gates (or "sluices"), rain is coming down. When it is dry and agreeable, he keeps the gates shut. In common language use, *Petrus* is referred to as the weather god, although, of course, there is no such thing in monotheist Christian belief prevalent in German-speaking regions. Why *Petrus* has come to mean the weather god only in German-speaking regions and apparently nowhere else in Europe remains a mystery. Whatever the reason for this development, *Petrus* as weather god has made his way from folk belief into everyday language use. As opposed to *Frau Holle*, who is mostly associated with snow events, *Petrus* is responsible for the full range of weather, as is evident in (283) and (284). However, people seem to reference him in particular where rain and/or good, or at least dry weather are concerned. The following example also lists *Graupel* 'graupel', which is a type of frozen precipitation that occurs in showers. It is opaque and softer than hail, but harder than sleet.

(283) ***Regen, Graupelschauer, Schnee, Sturm und eine dicke Nebeldecke*** – ***Petrus*** *hat jede Stunde wieder eine andere Idee.*
'**Rain, graupel showers, snow, storm and a thick blanket of fog** –**Peter** has a different idea every hour.'

(284) *Auch in diesem Jahr* ***hielt Petrus*** *für Tester und Produkte* ***einen*** *bunten* ***Mix aus Sonne, Regen, Hagel, Schnee und stürmischem Wind bereit****.* (*DWDS*)
'This year, too, **Peter had a** colorful **mix of sun, rain, hail, snow and stormy wind** for testers and products at the **ready**.'

An utterance from the *DWDS* historical corpus, given in (285), proves that *Petrus* was already in charge of letting it rain at least 150 years ago. The description of *Petrus* as treating the speaker 'heartlessy' indicates that the rain was not just unwelcome, but so plentiful that the speaker and his companions were soaked.

(285) *Gestern als* ***Petrus so herzlos mit uns umging, bis der Regen aus unsern Rockärmeln lief****, waren meine Hoffnungen auf den lange ersehnten Besuch des Bernina auf Null reducirt* [sic] … (*DWDS* : Swiss Alpine Club 1874)
'Yesterday, when **Peter treated us so heartlessly until the rain ran out of our coat sleeves**, my hopes for the long-awaited visit to Bernina were reduced to zero…'

Example (286) involves the blog entry of a tennis club and a hike the women's team had planned. No part of the blog entry contains any other mention of weather than the reference to

Petrus and that he 'did not mean well'. The reader does not learn whether it rained or snowed or whether it was just cold and windy, but the implication is that it was at least a rainy day.

(286) ***Petrus hatte es*** *allerdings* ***nicht so gut mit uns gemeint****, wovon wir uns unsere gute Laune aber nicht verderben ließen*. (*DWDS*: tc-vimbuch.de/ladies-morning-cup.de)
'However, **Peter didn't mean well (with us)**, but we didn't let that spoil our good mood.'

As already alluded to, "keeping it dry" is obviously the reverse of letting it rain and, thus, the suggested assertion of *Petrus* "meaning well", as in (287), implies there are no rain clouds in sight.

(287) *Aber* ***Petrus hat es*** *für heute offensichtlich* ***gut gemeint****, denn* ***ein blauer Himmel****, an dem nur ein paar vereinzelte* ***Schönwetterwolken*** *stehen,* ***verspricht einen strahlenden Tag****.* (roland-mueller-reiseberichte.de 'Roland Müller Travelogues')
'But **Peter** obviously **meant well** for today, because **a blue sky** with only a few isolated **fair-weather clouds promises a bright day**.'

Another frequent expression is that *Petrus* has *ein Einsehen* lit. 'an understanding' or 'sympathy' with the participants of any respective outdoor event in question when there is no rain. In a similar manner to (281) and (286), an explicit description of the weather conditions is not necessary. Example (288) only makes metaphorical mention of it not raining ('kept the floodgates closed'). The reader does not learn whether the sun is shining or whether it is cloudy. In this vein, the utterance is akin to any of the negated canonical MEs, predicate or argument type, such as *es regnete nicht* 'it did not rain' or *es fiel kein Regen* 'it fell no rain'.

(288) ***Petrus hatte*** *diesmal* ***Einsehen*** *mit den Radlern und* ***hielt die Schleusen dicht***. (DWDS)
'This time, **Peter was understanding** with the cyclists and **kept the floodgates closed**.'

The example in (289), too, contains the fixed expression *Einsehen haben* 'have an understanding' and makes no explicit mention of rain either. The metadata reveal that the utterance involves a bicycle race or trip. Even without further context, it is possible for the reader to infer that it was raining but then stopped after 20 minutes. Thus, (289) does not only assert the absence of rain from a certain point of time onwards, but it, in a way, contains a quasi-auxiliary: *Petrus* closes the gates and the rain *stops*. The speaker could have easily and in a more straightforward manner replaced the sentence with *nach 20 Minuten hörte es auf zu regnen* 'it stopped raining after 20 minutes' instead of this metaphorical circumlocution.

(289) *Nach 20 Minuten war es endlich soweit,* ***Petrus*** ***hatte ein Einsehen und*** ***schloss die Pforten****.*
(*DWDS*: fittyundbussiontour.com)
'After 20 minutes it finally happened, **Peter had sympathy and closed the gates**.'

Although there are several fixed expressions that collocate with *Petrus* (*Einsehen haben* 'have understanding', *gnädig sein* 'be merciful', *es gut/nicht gut meinen* 'mean/not mean well'), there is no limitation really to how *Petrus* proceeds with precipitation. His dealings with weather events are often achieved by describing his "frame of mind", i.e., his benevolence or pitilessness, given in one clause followed by the ME in another clause. These clauses are at least "heavy+heavy" when it comes to the semantic load, but they can also be "triple heavy", as in (289) (Peter+having sympathy+closing the gates). In clauses, which express that *Petrus* ***ließ*** *es regnen /* ***ließ*** *Regen fallen / schickte Regen* 'Peter let it rain / let rain fall / sent rain', there are "light" elements, such as causal *lassen* 'let'. Obviously, in the first two examples, canonical constructions are merged with a transitive causal agent so that there are each two heavy and two light elements involved.

It is not necessary, however, to employ one of the more canonical ME constructions. Any expression that contains an ME can contextualize the ME in addition to referring to *Petrus.* This, furthermore, allows the speaker to project their own stance regarding the ME onto *Petrus*, as the responsible entity. In (290) and (291), there is no causal connection between *Petrus* and the precipitation events. *Petrus* does not 'send' rain or 'let it rain'. The reader must make the connection based on pragmatic inference. And the speaker assumes that the metaphor of *Petrus* is strong enough for the reader to associate *Petrus* with the respective ME.

(290) ***Petrus war gnädig*** *mit uns und* ***es kam kein Regen***. (*DWDS*: co-natur.de)
'**Peter was merciful** with us and **it came no rain**.'

(291) ***Mit Schnee war*** *ja* ***nicht viel*** *– wofür* ***ich Petrus danke****.* (*DWDS*: wiggerlsworld.wordpress.com)
Lit. '**With snow wasn't much** – for which **I thank Peter**.'

The same is true for (292) where the precipitate *Regen* 'rain' is mentioned, but the fact that 'Peter watered the lawn', in the context of a soccer game, could just as well lead readers unfamiliar with *Petrus* to believe he was a groundkeeper since soccer pitches routinely get watered before a game and during halftime in order to increase the speed of the ball.

(292) *Das Wetter machte am Samstag glücklicherweise auch mit, es war nicht zu kalt und* ***auch mit dem Regen hatte Petrus ein Einsehen und wässerte*** *nur in der Halbzeitpause* ***kurz den Rasen****.* (fufuswelt.de 'Fufu's world')
'Fortunately, the weather cooperated on Saturday, it wasn't too cold, and **Peter also was understanding regarding the rain and only briefly watered the lawn** during the half-time break.'

Lastly, in (293) *Petrus* is not just the source or causer of the rain, he does not only send it or arrange for it, but he literally "manufactures" the precipitate.

(293) *Während der Hinfahrt* ***machte Petrus Regen****,* ***viel Regen****.* (*DWDS*, 3-paws.de)
'During the journey there **Peter made rain**, **lots of rain**.'

The utterance above could be rephrased in several ways including more canonical constructions, such as *während der Hinfahrt regnete es in Strömen* 'it was pouring (lit. raining in streams) during the journey' or es gab *strömendem Regen auf der Hinfahrt* 'it gave pouring rain on the journey'. In both of these paraphrases, the ME is profiled. Another way to profile the event in a main clause is to use a "pseudo-progressive" with a preposition, such as *auf der Hinfahrt war es am Regnen wie*… 'on the journey it was raining like….'. Or, the ME can also be backgrounded in another paraphrase, which employs a preposition but no predicate: *bei strömendem Regen fuhren wir nach* … 'at streaming rain we drove to…'. The next section concerns the last two options and explores the use of the *am*- and *beim*-progressives as well as the predicateless use of prepositions, most notably *bei* 'at', in combination with an ME.

4.3.2 Prepositional constructions

This section explores two minor constructional schemas used for MEs, which place the precipitate as a prepositional complement within a prepositional phrase, rather than in a predicate-argument construction. The first of these preposition-based constructions are actually examples of the "pseudo-progressive" construction in German or extensions thereof, namely *es ist am Regnen* 'it is at (the) raining' and *es ist beim Regen* 'it is at (the) raining', both of which more felicitously translate to 'it is raining'. The second of these preposition-based constructions involve a nominal form of the precipitate, in fact, a bare nominal form with no determiner, such as *bei Regen* 'at rain'. Rather than signaling progressive force, these latter constructions suggest that the ME co-occurs with the event profiled in the main clause. I discuss each of these broad classes of prepositional ME constructions in turn.

Since German, unlike English, lacks overt grammatical aspect and speakers are not linguistically obligated to express progressive aspect, utterances with progressive force are usually given in the simple present (*Präsens*). However, German speakers have several linguistic means at their disposal to express that an action is ongoing at the moment of speaking or ongoing at the time another event is taking place. They may opt for one of these linguistic means for emphasis or clarification, or if the situation requires it. One option is to add the adverb *gerade* (lit. 'straight', roughly 'just' or '[right] now') after the verb, as in (294).[136]

(294) *Ich lese **gerade***.
'I am reading.' (lit. 'I read **right now**')

The other option for signalling progressive force is the use of the just mentioned periphrastic constructions such as the so-called "*am*-progressive" or the "*beim*-progressive", as in (295).[137]

(295) *Ich* *bin* *am/beim* *Lesen.*
I be.1SG.PRES PREP+the.DAT read.INF
'I am at the reading.'

These constructions are formed by using a form of *sein* 'be' and either the preposition *an* 'at' or *bei* 'near' with an infinitival (effectively, a nominalized) verbal complement. These infinitival complements are considered to be nominalized since they provoke the addition and ultimately fusion of the neuter definite determiner *dem* in the dative case with the preposition to form *am* or *beim*. It is also possible to combine both of these periphrastic progressive aspect strategies as in (296).

(296) *Ich bin **gerade am/beim Lesen.***
'I am **at the reading right now**.' / 'I am reading.'

The original meaning of the preposition *bei* is akin to English *by* and signals nearness or vicinity of a reference landmark. As such, any complement of *bei* can assume a locative reading.

---

[136] To be more precise, *Präsens* (Simple Present Tense) and *Präteritum* (Simple Past) elicit a continuous aspect reading, whereas *Perfekt* (Present Perfect) generates an immediate anteriority reading (e.g., *Ich habe **gerade** einen Brief geschrieben* 'I just finished writing a letter') (cf. Schaden 2010).

[137] The origin of the "*am*-progressive" was originally attributed to West Germany (especially the area around Cologne) and came to be known as *rheinische Verlaufsform* **'**Rhenish progressive form', but studies from the 1990s onwards have shown that the *am*-progressive is now widely established in Germany, Austria, and Switzerland (Reimann 1998, Krause 2002, Elspaß and Möller 2003, Van Pottelberge 2004). Originally a regional dialect, its expansion has also helped the *am*-progressive to become more prevalent in standard written German (cf. Gárgyán 2014).

According to Spreng (2017), the *beim*-progressive is seemingly preferred for utterances in the past tense, which indicates that not just physical absence may play a role but also temporal absence.[138] There appears to be some disagreement over how the infinitive that follows the preposition should be analysed, which, ultimately, could influence how the type of construction can be labelled with regard to MEs. The majority of analyses hold that *am* and *beim* are contractions of a preposition and a neuter definite determiner in dative case (e.g., Krause 2002, Delisle 1986, Donaldson 2007, Ebert 2000). Correspondingly, the infinitive that follows must be nominalised as the complement of the preposition. Spreng (2017) claims that the *am*-progressive needs to be analysed as an aspectual particle plus a complex verbal construction or a VP complement. These matters cannot be resolved in this dissertation, but for the sake of the analysis at hand it must suffice to subsume these constructions as [PREP + $PRED_{prog}$ + ME] and as [*bei* + ME_{ARG}]. The analysis of the construction furthermore influences the orthography. In German, nouns are usually capitalized and, in accord with the traditional analysis, *Duden* gives the *am*+V.*sein* 'be at V' examples with capitalized verbal nouns: *das Wasser ist am Kochen* 'the water is at boiling', *ich bin noch am Überlegen* 'I am still at thinking', *er ist am Essen* 'he is at eating'. Reseach papers on the German progressive as well as speakers on the web differ in their choice of using upper- or lowercase, which likely mirrors the uncertainty of the status of the construction: is the precipitate a verbal infinitive or a de-verbal nominalization? The following examples all reproduce the speaker's choice of orthography. I will use *r/Regnen* and *s/Schneien* etc. to capture results that include both upper- and lowercase spellings of the nominalized infinitival verb. All examples in this dissertation are reproduced as written by the speaker.

Table 102 shows the frequencies for the *beim*- and *am*-progressives in combination with the four precipitate types *Regen* 'rain', *Schnee* 'snow', *Schneeregen* 'sleet', and *Hagel* 'hail'. The *beim*-progessives show low incidences for each precipitate. In fact, there are no returns for *beim s/Schneeregnen* 'at sleeting' and *beim h/Hageln* 'at hailing'. The *am*-progessive occurs with the highest incidence for the most frequently occurring precipitate, *Regen* 'rain' (N=813). It occurs with the second highest incidence for *Schnee* 'snow' (N=167) in both corpora, the *deTenTen13*

[138] Spreng exemplifies that a *beim*-progressive as in *Er ist beim Schwimmen* 'he is at/by-the swimming' suggests that the event takes place away from the speaker, as in 'he is *off* swimming', while the *am*-progressive *Er ist am Schwimmen* 'he is at-the swimming' does not trigger this reading (2017: 1). Take note that Spreng's focus lies on the Swabian dialect specifically, which may differ from other regional uses of the two progressives.

and the *DWDS*.[139] *Schneeregen* 'sleet' and *Hagel* 'hail' together only return a handful of results from either corpus.

Table 102. *Frequencies of the* am- *and* beim-*progressives in the* deTenTen13 *and the* DWDS Metakorpora

Precipitate	***deTenTen13 am* 'at the'**	***DWDS am* 'at the'**	***deTenTen13 beim* 'at the'**	***DWDS beim* 'at the'**
r/Regnen 'raining'	813	315	28	35
s/Schneien 'snowing'	167	42	9	30
s/Schneeregnen 'sleeting'	2	0	0	0
h/Hageln 'hailing'	3	3	0	0
Total	**985**	**360**	**37**	**65**

Miller et al. (2017) found that German L1 speakers "make a significant and crucial distinction between conditions in which *beim* is grammatical and others in which it is unambiguously ungrammatical. Thus, although they show a general preference for *am* even where both *am* and *beim* are acceptable, their usage of *beim* is not random" (2017: 224). Here, Miller et al. (ibid) give the example **beim Regnen*, specifically. This finding seems to be in accordance with a number of *Google* results that I came across in which learners of German asked the Web community for clarification on the difference between the *am*- vs *beim*- progressive and how or when to use each correctly.[140] Barrie and Spreng (2009) claim that *beim* is more colloquial than *am* depending on the German dialect and a speaker's own preferences and grammaticality judgments. At any rate, based on my own judgement as a native speaker, some of the returned results (especially from *Google*) sound clumsy or very informal at best, as in (297).

(297) *Im Frühjahr **beim Regnen** ist die Cam komplett mit Wasser vollgelaufen.* (wald-prinz.de 'forest-prince')
'In spring, **at the raining**, the camera got wet (lit. completely filled up with water).'

We cannot exclude the possibility that *beim*-utterances are the result of a typo where the author wrote *Regnen* (infinitival verb 'to rain') instead of *Regen* (N. 'rain'), but they may also simply be the result of a regiolect or personal preference. A more standard version would be *als*

[139] The *DWDS* itself contains a number of corpora, which themselves are subsumed in *Metakorpora* depending on their nature (e.g., web vs. newspaper) and time frames. The respective metacorpora are the *Gegenwartskorpora mit freiem Zugang* 'contemporary corpora with free access', *Historische Korpora* 'historical corpora', *DTA-Kern+Erweiterungen* 'DTA-core+extensions', *ZDL-Regionalkorpus (ab 1993*) 'ZDL-regional corpus (from 1993)', which contains newspaper articles, and the *WebXL* corpus. The corpora that are part of each metacorpus can be found here: www.dwds.de/d/k-meta (last accessed April 10, 2024).
[140] For example, *UniLang Language Community*, *HiNative.com*, *tr-ex.me (Translate Examples)*, etc.

es regnete 'when it was raining' for a verbal construction and *bei/m Regen* 'at rain' for a preposition+noun construction rather than *beim Regnen* 'at the raining'. In example (298), the meaning of *beim Regnen* is a more generic one, rather than progressive or continuous, as in *when(ever) it rains, these types of plants clean themselves*. It should be noted that *bei Regen* 'at rain' would appear to be the more felicitous choice.

(298) *Am Beispiel der Natur können wir aber lernen, wie eine Selbstreinigung aussieht: Kapuzinerkresse, Kohl, Lotus etc. reinigen sich* ***beim Regnen*** *selbst.* (durexon.de)
'Using nature as an example, however, we can learn what self-cleaning looks like: nasturtiums, cabbage, lotus, etc. clean themselves **at the raining**.

Many of the *beim* examples collocate with main verbs of perception, i.e., *zuschauen*/*zugucken* 'to watch' or *zuhören* 'listen to', as shown in examples (299) and (300). These examples, furthermore, echo the cognate ME constructions.

(299) *So lief ich also im Regen zur Hotellobby, wo ich warm und trocken im Netz surfen kann und* ***sehe*** *hier seit Stunden* ***dem Regen beim regnen*** *zu.* (kulturweit-blog.de)
'So I walked in the rain to the hotel lobby, where I can surf the Net warm and dry and have been watching the rain fall for hours/lit. **watching the rain at the raining**.'

(300) *Nachdem ich heute Vormittag gearbeitet, meine Einkäufe erledigt und Kuchen gebacken habe, hüpf' ich jetzt gleich aufs Sofa für ein gemütliches Strickstündchen und* ***schau' dabei dem rieselnden Schnee beim Schneien zu****.* (andans.de)
'After I worked this morning, did my shopping, and baked a cake, I'll hop right onto the sofa for a cosy hour of knitting and **watching the trickling snow at the snowing**.'

Interestingly, a number of results that *Google* returned were from publications of the 19th century, including German textbooks that contained *beim Regnen* examples. The example in (301) is from a textbook written for German teachers. The example in (302) is from a textbook for teachers for the deaf.

(301) *Warum kannst du die Sonne* ***beim Regnen*** *nicht sehen? Weil sie über den Wolken ist.* (G. Heneß 1884: 100)
'Why can you not see the sun **at the raining**? Because it is above the clouds.'

(302) *Wenn wir einen Regenschirm* ***beim Regnen*** *über uns halten, so hält derselbe den Regen von uns ab*. (Wirsel 1854: 212)
'If we hold an umbrella above us **at the raining**, it keeps the rain off us.'

In these textbooks, neither example is part of teaching conversation about the weather, but part of definitions of the verbs *verschwinden* 'disappear' and (*ab*)*halten*, roughly 'hold/halt/hinder'. In both cases, the constructions do not encode progressivity *per se*, but conditionality. Example

(303), however, refers to two continuous events in the past and reads as 'when (we) were crossing the Red Sea and when it was raining manna'.

(303) *… das Wunder beim Durchqueren durch das rote Meer und **beim Regnen** von Manna.*
(S. A. Poznański 1901: 748)
'… the miracle at crossing the Red Sea and **at the raining** of manna.'

It is always intriguing when speakers use different constructions in one and the same text for either the same precipitation event or for different precipitation events that happen at the same time. The speaker in (304) uses three different constructions for talking about something that always happens to his car when it is raining. Seeking advice in a car forum, he writes that he noticed that his car's indicators emit sparks when he is changing lanes and driving at a certain speed while it is raining. First, he employs two predicate constructions. He then uses *beim regnen* 'at the raining' followed by *bei Regen* 'at rain'.

(304) ***Es hatte unendlich geregnet** und ich fuhr mit knapp 110 km/h.* (…) *Das ist aber nur so **wenn es stark regnet.** Bei unter 80 **beim regnen** sehe ich nichts! Habe ihn letztes mal abgestellt und **bei Regen** draußen mal angeguckt, aber nichts zu sehen!* (ford-forum.de)
'**It had rained endlessly,** and I was driving at almost 110 km/h. (...) But that's only the case **when it rains heavily**. If it's under 80 **at the raining**, I can't see anything! I parked it (the car) last time and **at rain** looked outside, but nothing to see!'

In essence, the speaker could have employed the predicate construction four times. The *beim regnen* construction seems clumsy to me and could effectively have been more felicitously substituted with another *bei Regen* 'at rain'. It is possible he only used it to avoid writing *wenn es regnet* twice in a row. In fact, the German language learning blog *Deutsche Grammatik 2.0*, run by a linguist, has a nominalisation exercise called *Sonne und Regen* 'sun and rain' where the learners are asked to transform *wenn es regnet* 'when it is raining' into *bei Regen* 'at rain'. *Bei Regen* appears to be the most elegant linguistic solution, but before I discuss this construction, let us turn to the use of the second "pseudo"-progressive, i.e., the *am*-progressive, which is deemed the more acceptable of the two, i.e., *beim* vs. *am*.

Gárgyán (2014) identifies eight different uses of the *am*-progressive, but the majority of authors seem to agree on the *am*-progressive as expressing an internal perspective that is used to encode an event or activity as ongoing. As opposed to the *beim*-progressive, as discussed above, the *am*-progressive lacks the force of a locative reading. In short, there is semantic overlap between *am* and *bei*, but while the first one, according to Ebert (1996) and Miller et al.

(2017), requires the subject to be where the activity is taking place, the latter only requires the subject to be engaged in the ongoing activity. For example, a speaker who is *beim Schwimmen* could possibly be a parent watching their child at the swimming pool/at swimming practice, but they not necessarily have to be engaged in swimming themselves. A speaker who claims to be *am Schwimmen* must be actively engaged in the activity of swimming (cf. fn.138). Despite the fact that the speaker herself is not actually engaged in an ME, but merely affected by it, the *am*-progressive does seem more immediate in its construal. By that I mean that the ME more directly impacts the speaker. This is presumably due to the *es* 'it' actively carrying out the activity of "weathering"—as diffuse as *es* 'it' may be—while the *bei*-construction backgrounds the ME. The *beim*-progressive, too, appears more circumstantial or peripheral. Consider examples (305)–(308), where each example contains a different precipitate type. Here the uncertainty regarding capitalization is obvious with the first three MEs (including *am stürmen* 'at the storming' in (306)) written in lower case and the last two capitalized.

(305) *Der Winter kommt nun doch. Seit circa 5 Stunden* ***ist es*** *hier* ***am schneien****. (island-of-hope.de)*
'Winter is coming after all. **It has been at the snowing** here for about 5 hours.'

(306) *Die* ganze *Fahrt über und auch jetzt noch* ***ist es am stürmen*** *und teilweise heftig* ***am regnen****.* (didi-kiel.de 'Didi Kiel')
'The whole trip and even now **it is** still at the storming and to some extent **at the raining** heavily.'

(307) *Nur die Autofahrer bekamen nicht mit, dass* ***es*** *derbe* ***am Schneeregnen war****, und dass man nicht in einem Waldgebiet mit Wildwechsel auf einer Schlaglochpiste mit 80 km/h fährt, wenn nur 50 erlaubt sind.* (birgit-edler.de)
'Only the car drivers didn't realize that **it was** really **at the sleeting** and that you don't drive at 80 km/h on a pothole road in a wooded area with deer crossing when only 50 km/h are allowed.'

(308) *Krass. An der Mosel bei den Eltern einer Bekannten* ***war es*** *wohl* ***so am Hageln,*** *dass der Wohnwagen, den sie da haben, und das Auto total zerdellt oder teils sogar kaputt gegangen sind. Scheiben alle im Arsch, Motorhaube und Dach (…) zerdellt...ganz übel…* (bekloppt-unberechenbar.joyclub.de 'bonkers-unpredictable.joyclub.de')
'Gnarly. On the Moselle, at the parents of a friend of mine, **it was** (reportedly) **so at the hailing** that the caravan they have there and the car were totally dented or even partially broken. Windows all broken (vulg.), hood and roof (…) dented… really nasty...'

With the exception of (305), these examples include modifiers of MANNER *heftig* 'severely', *derbe* 'mightily/really' (colloq.), and *so* 'so' and appear very emphatic. Furthermore, the first two examples describe the situation at the point of writing, while the last two examples describe past events. Still, there is a palpable immediacy to (307) and (308), which is construed by the use of the *am*-progressive rather than the use of a predicate construction.

An example that offers insight regarding the different construals of the predicate construction and the *am*-progressive can be found in (309). Indisputably, the speaker is aware that there is a difference in the constructions that she juxtaposes. The speaker talks about riding her pony in the rain the previous day where her pony did not cough (for once), which she believes was due to the rain preventing dust. The first construction (past perfect: *gestern hatte es geregnet*) could mean that it rained the previous day, but not necessarily that the speaker was out riding while it was raining. Maybe she went for a ride after the rain stopped. The second construction (involving the *am*-progressive: *bzw. war am regnen)* implies that the two events took place at the same time, i.e., it was raining while she was riding her pony. The concordance alone does not necessarily spell out the events, but the wider context of her blog entry does. Furthermore, the abbreviation *bzw*. stands for *beziehungsweise* and can be translated either as 'respectively' or 'rather'. The latter, meaning 'rather' or 'more precisely', is more common in informal language (cf. *DWDS*) and the preferable reading respecting the utterance.

(309) *und viel besser: keinmal gehustet.. denke doch das es am staub lag.. denn* ***gestern hatte es geregnet bzw war am regnen****.* (fjordblog.de)
'and much better: no coughing… (I) think that it was the dust … because **yesterday it had rained respectively was at raining**.'

Lastly, the *beim*-data revealed one construction that has not yet been mentioned but frequently appears in description of weather events: it is that of a prepositional phrase, i.e., a preposition plus a precipitate in its argument form. Most notably, this concerns the preposition *bei* 'at/by', which is related to and has already been explored as part of the *beim*-progressive. Since the [PREP+ARG] construction does not contain a predicate, it cannot stand on its own, but it is ordinarily used in lieu of a subordinate clause. More specifically, it serves as a conditional clause, which usually expresses an action or outcome that occurs if a certain condition is fulfilled first. In German, these conditional sentences are typically introduced by subordinating conjunctions *wenn* 'when/if' and *falls* 'if'. This is a very condensed grammatical description. Among others, another possible—but more rare—subordinating conjunction introducing a conditional is *sofern* 'if/provided that'. Both *falls* 'when/if' and *sofern* are more elevated in style than *wenn* 'when'. Furthermore, it is possible to move the verb to the sentence initial position and eliminate *wenn*, but this construction is seldom used in spoken language (e.g, *Ist das Wetter gut, gehen wir spazieren*. Lit. 'Is the weather good, go we for a walk.' or *Hört der Regen auf, gehen wir spazieren*. Lit. 'Stops the rain, go we for a walk.'). The utterance in (310) contains this

type of conditional clause and example (311) demonstrate how this type of conditional can be restated with a prepositional phrase involving *bei* 'at' where it is also possible to change the word order to *Wir machen **bei gutem Wetter** einen Spaziergang*. Lit. 'We make **at/by good weather** a walk.'

(310) ***Wenn das Wetter gut ist**, machen wir einen Spaziergang.*
'**When/If the weather is good**, we will go for a walk.'

(311) ***Bei gutem Wetter** machen wir einen Spaziergang.*
Lit. '**At/by good weather**, we go for a walk.'

In both instances, the weather is not the main event but a circumstantial event that nevertheless affects the main event. Correspondingly, the ME in question is backgrounded, in contrast to the *am*- and *beim*-progressives, which profile the ME.

I carried out a simple concordance search for *bei* 'at' and the respective precipitate in the *deTenTen13*. The frequencies are given in Table 103. The frequencies mirror other results in that *Regen* 'rain', as the most common precipitate, has the highest incidence with almost four times as many returns as *Schnee* 'snow', which has the second highest incidence. *Hagel* 'hail' slightly outweighs *Schneeregen* 'sleet', but both precipitates have considerably lower frequencies than *Regen* 'rain' and *Schnee* 'snow'. The returns from the *DWDS* metacorpora are even more frequent than those from the *deTenTen13*, especially with regard to *Regen* 'rain' or *bei Regen* 'at rain'.

***Table 103.** Frequencies of preposition* bei *'at' plus noun precipitates in the* deTenTen13 *and* DWDS Metakorpora

Precipitate	*deTenTen13*	*DWDS*
***Regen** 'rain'*	45,569	129,655
***Schnee** 'snow'*	12,380	20,472
***Schneeregen** 'sleet'*	381	795
***Hagel** 'hail'*	437	742
Total	58,767	151,674

Regen 'rain' is the most frequently occurring precipitate type, and it is also typically the most long-lasting ME of the four. Take note that according to *wetter.de* 'weather.de', a bout of hail usually does not usually last longer than 15 minutes, which is shorter than most activities that people carry out "at" some type of weather, especially if it is outside, like a run or a hike. Therefore, it seems straightforward that *bei Regen* 'at rain' has the highest frequency in this construction.

Table 104 lists the definitions and usage conditions of *bei* 'at/by' as given by the *DWDS* dictionary and as they apply to the [PREP+ARG] construction. The meanings, usage conditions, and examples were checked and counterchecked in three sources, the *Duden online*, the *DWDS*, and *Hammer's German Grammar* (Durrell 2011). The examples for the different meanings and usage conditions happened to include a weather example for each definition in question, therefore I adopted the version from the *DWDS* as given in Table 104. It should be noted that there appears to be great overlap between the "types". I would argue that all types designate the simultaneity of the respective ME with another event and that (iii)–(v) are more specific categorizations of (i) and (ii). Accordingly, I do not differentiate between (i) and (ii).

Table 104. *Definitions and examples of the preposition* bei *'at/by' adopted from the* DWDS *dictionary*[141]

	Types / definition	German example	English translation
i.	**designates the simultaneity of a state or process with another state or process (occurring at the same place)**	***bei dem*** *gestrigen* ***Regen*** *fanden wir unter einem Baum Schutz*	'we found shelter under a tree **when it was raining** yesterday' (lit. '**at the** yesterday's **rain**')
ii.	**attendant circumstance**	*der Zug passierte die Grenze* ***bei Regen***	'the train passed the border **at rain**'
iii.	**conditional**	***bei Regen*** *gehen wir nicht*	'**at rain** we don't go for a
iv.	**causal**	***bei dem Regen*** *sind die Wege unpassierbar*	'**at the/this rain**, the roads are impassable'
v.	**concession**	***bei so starkem Regen*** *sind sie spazierengegangen*	'**at such hard rain** they went for a walk'

Despite some overlap or unclarity, the *DWDS* definitions above have already provided an excellent introduction into the different uses of *bei+Regen* 'at rain' with very natural examples, where the construction acts as a stand-in for some type of subordinating phrase (conditional, concessive, causal, etc.). In examples (312) and (313), the topic of the main clause is generally true if the real or possible condition of rain is given.

(312) *… man kann ja nicht die ganze Zeit im Gastzimmerchen hocken. Und* ***bei Regen*** *****ist man auch nicht gern lang am Strand.*** (community4um.de)
'… you cannot sit in your little guestroom all the time. And **at rain**, **you do not like to spend a long time on the beach either**.'

[141] Definitions and all German examples are taken from the *DWDS* (based on the *eWDG*), translations are mine. Source: www.dwds.de/wb/bei#1 (accessed January 20, 2022).

(313) *Bisher wusste ich gar nicht, was ein Wetterpilz ist. Jetzt weiß ich es auch:* ***Ein pilzförmigs Schutzdach (…) unter das man sich bei Regen unterstellen kann****.* (taf-timing.de)
'Until now, I didn't know what a "weather mushroom" is. Now, I know: **a mushroom-shaped shelter (…) under which you can seek cover at rain**.'

Example (314) contains a continuation of variable ME constructions involving rain, wind, and sun. The focus of the dissertation at hand is on precipitation, but other MEs, especially when they appear in the same utterance, are still compelling and behave in similar ways.

(314) *(Die Reise einer unbekannten Person) Sie startete* ***bei Regen*** *in Hamburg, machte einen Zwischenstopp* ***bei Wind*** *in München und landete schließlich* ***im sonnigen Rom****.* (magischerfc.de – magicalfc.de)
'(The journey of an unknown person) She started **at rain** in Hamburg, made a stopover **at wind** in Munich and finally landed **in sunny Rome**.'

In (314) the traveler starts *bei Regen* 'at rain' in Hamburg, which in English can be translated with a preposition as well: 'in the rain', or with a more complex relative clause 'where it was raining'. Although the second ME employs the same construction (*bei Wind*), a translation with a preposition 'in the wind' takes on a different, more idiomatic meaning in English. *In wind* alone sounds odd, although in German—as in English—there is the idiom *bei Wind und Wetter* 'in [any] wind and weather'.[142] Another paraphrase with a cumbersome relative clause is also a possibility: 'where it was windy' or 'where there was wind (blowing)', but obviously these also do not parallel 'where it was (at) raining' (*'*where it was [at] winding'*). The last ME *im sonnigen Rom* can be directly translated into English as 'in sunny Rome'.

In Chapter 3, I pointed out that for many of the *Schnee* 'snow' expressions, it is difficult to tease apart whether the utterance involves the actualized ME—snow that is falling—or snow that is merely lying on the ground. In (315) and (316), it is clear that the ME involves snow that is in the process of falling. In (315), it is obvious that one would only undertake a snowshoe tour when there is already a thick layer of snow on the ground. Consequently, spelling out that they did the trek 'at snow' must refer to snow falling while the speaker is simultaneously plodding

[142] A search query in the *deTenTen13* for *bei jedem Wind und Wetter* 'in any wind and weather' returned 432 instances. The *deTenTen13* returns 22,660 concordance lines in total for a simple *bei Wind* 'at wind' search query. A total of 12,600 of those concordances are the idiomatic *bei Wind und Wetter* so that the remaining 10,060 concordance lines do not include *Wetter.* Many of these utterances enumerate other MEs. Where *bei Wind* appears by itself, it is often with more typical "wind-related" activities, such as flying a kite, being at the sea, or riding a bike, rather than merely "arriving" somewhere where it is windy, as in (314).

through a layer of snow. The second clause remarks on a second tour, which took place 'at picture-book weather'; presumably, this involved blue skies and sun with no precipitation falling. The term *Bilderbuchwetter* if further explored in the next section, which concerns weather compounds with *Wetter* 'weather'.

(315) *Auch zwei herrliche Schneeschuhtouren, einmal **bei Schnee und heftigem Wind** und einmal **bei Bilderbuchwetter** konnte ich noch machen*. (laufen-martin.blogspot.ch 'running-martin.blogspot')
'I was also able to do two wonderful snowshoe tours, one **at snow and strong wind** and one **at picture-book weather**.'

In (316) it is specified how, especially when snow is falling, the light on the little snowflakes provides a delightful picture. Again, it is unlikely that the speaker refers to single snowflakes in a mass on the ground. Besides, the lighting towers (former GDR watchtowers) always produced a special light, and so it is only under the condition of snow that the lighting appears especially enchanting.

(316) *Das waren ganz besondere Beleuchtungstürme, die (…) insbesondere **bei Schnee** eine wunderbare Strahlung auf die kleinen Schneeflocken zauberten.* (wanderboje.de 'wandering-buoy')
'These were very special lighting towers that (…) conjured up a wonderful light on the small snowflakes, especially **at snow**.'

Lastly, regarding precipitation types *Schneeregen* 'sleet' and *Hagel* 'hail', the latter is much more likely to occur as the conditional type, in particular with measures taken and insurance carried out in anticipation of hail damage. Still, the corpus searches returned enough examples where the circumstances are not conditional but attendant, as in (317), which reports on a bicycle race. Obviously, the weather conditions were difficult, and the race was accompanied and hampered by bouts of hail and downpour. It seems likely that the cyclist crossed the finish line while it was hailing.

(317) *Timo Schäfer wird **bei Hagel und Platzregen**, in einer der schwersten Meisterschaften der letzten Jahre, Vize-Saarlandmeister.* (bike-aid.de)
*'**At hail and downpour**, Timo Schäfer becomes runner-up in Saarland in one of the most difficult championships in recent years.'*

The example in (318) involves *Schneeregen* 'sleet', which is presumably less likely to occur as the conditional type. The use of the adverb *gerade* 'right now' clearly signals progressive aspect and the attendant circumstance or simultaneity of the author sitting and writing his "comments" while the sleet event provides the background.

(318) *Unser gemeinsamer Andalusienurlaub ist jetzt leider schon vorbei und ich sitze gerade **bei Schneeregen** und gebe meinen Kommentar ab.* (siggis-hobbyradler.de 'siggis-hobbycyclists')
'Unfortunately, our vacation together in Andalusia is now over and I'm just sitting **at sleet** making my comment.'

Following the definitions provided by the *DWDS*, I suspect that many of the *bei Hagel* 'at hail' examples are conditional or causal rather than "just" attendant and that they involve protection from possible hail events, such as noted in insurance terms or as commented on in (319).

(319) *Jetzt hab ich ein tolles Carport und bin so froh das ich mich dafür entschieden habe. Kein Schnee vom Auto machen mehr, geschützt vor z.B. kleinen Ästen die runterfallen, und vor allem **bei Hagel**. (finanzratgeber24.de 'financial-advisor24')*
'Now I have a great carport and I'm so glad I decided for it. No more snow (to brush) from the car, protected from, for example, small branches that fall down, and especially **at hail**.'

In order to confirm my suspicions regarding the usage of conditional meaning (can the utterance be rephrased with *if/when*) vs. simultaneity (does or did the ME take place concomitantly with another, profiled event), I looked at the returns for each precipitate with *bei* 'at'. I randomized the results using the shuffle function and then determined the usage contexts for each of 100 concordance lines per precipitate type in total. In order to randomly select returns for further scrutiny, I examined the usage for every third concordance line on every 10th page for the *Regen* 'rain' and *Schnee* 'snow' precipitates and for every second concordance line on every second page for the less frequent precipitates *Hagel* 'hail' and *Schneeregen* 'sleet'. The frequencies are given in Table 105. Take note that the majority of the returned concordance lines for the *Schnee* 'snow' precipitate license the reading of either snow on the ground or snow falling, or indeed both at the same time.

Table 105. *Frequencies of* bei*+precipitate (conditional and simultaneous events) out of 100 concordances per precipitate in the* deTenTen13 *corpus*

Usage type / Precipitate	**Conditionality** ('if/when')	**Simultaneity** ('while')
Regen 'rain'	79	21
Schnee 'snow'	90	10
Schneeregen 'sleet'	23	77
Hagel 'hail'	88	12
Total	**280**	**120**

The precipitates *Regen* 'rain', *Schnee* 'snow', and *Hagel* 'hail' chiefly occur in utterances that express conditionality, as in examples (312), (316), and (319). Only the *Schneeregen* 'sleet' precipitate occurs predominantly in utterances expressing simultaneity, as in (318). Again, these incidences seem intuitive considering that *Schneeregen* 'sleet' is not a "basic" or canonical precipitate, like rain or snow, where either of the latter would be the expected precipitate type "in the case of …" or "on condition of …", as in (320). It would be odd, to make *Schneeregen* 'sleet' the "default attendant ME", as in an event taking place inside *bei Schneeregen* 'at sleet' or 'if it is sleeting', but not if it is raining.

(320) ***Bei Regen*** *findet die Veranstaltung im Saal statt*. (oegnet.de)
'**At rain**, the event will take place in the hall.'

Finally, the *bei*+ME examples revealed another trove of prepositional uses, which I shall only mention in passing, but which I hope to explore in further research. The prepositions in question are *in* 'in' or *im* 'in the' (in+dem 'in+the.DAT') and *durch* 'through'. In particular *im* 'in the' appears to overlap with *bei* 'at', but there are certainly restrictions. For example, *bei* 'at' can be an attendant circumstance, which does not necessarily have to affect the speaker physically, while *im* 'in the.DAT' implies that the person experiencing the ME is directly exposed to it. Consider example (321), in particular, where both *bei* 'at' and *im* 'in the' seem equally fitting.

(321) *Doch wenn das Wetter nicht mitspielt, endet das Vergnügen schnell.* ***Im Dauerregen***, ***bei Schnee*** ***oder Frost*** *nimmt die Zahl der Fußballer schnell ab* (Fussball-euro2008)
'But when the weather doesn't play along, the fun ends quickly. **In the continuous rain, at snow or frost**, the number of footballers decreases quickly.'

While *bei* 'at' seems indeed "attendant", *im* 'in the' construes the event as impacting the experiencers more strongly. People carry out all sorts of activities *bei* 'at' any sort of weather, but rarely do they do so *in* 'in' a certain type of weather. Accordingly, people drive their cars or go shopping *bei Regen* 'at rain' or *bei Schneeregen* 'at sleet', but this does not mean they literally experience the precipitate directly, such as getting wet, as in (322). It could mean that the speakers were en route to their destination in the car or some other vehicle without necessarily being exposed to the elements. In (323), it is obvious that the speaker was not able to shelter from the hail. He points out that not everything went smoothly.

(322) *Heute ging es* ***bei Schneeregen*** *nach Innsbruck, um ein paar Einkäufe zu erledigen.* (Sport10.at)
'Today (we) went to Innsbruck **at sleet** to do some shopping.'

(323) *Die Fahrt verlief bis auf einen Sprint in Wuppertal Barmen (zum Bus) reibungslos. In Sprockhövel angekommen, mussten wir erstmal 10min **im Hagel** zur Bildungsstätte laufen...* (marcus-in-hd.de)
'The journey went smoothly except for a sprint in Wuppertal Barmen (to the bus). At arrival in Sprockhövel, we had to walk 10 minutes **in the hail** to the educational center...'

In contrast to example (313), the prepositions *bei* 'at' and *im* 'in the' in examples (322) and (323) are interchangeable. It is possible to seek shelter *vor Regen* 'from rain' and *bei Regen* 'at rain' or 'when it is raining', but it seems rather odd to seek shelter *im Regen* 'in the rain'. The function of the preposition *in*, spatially, means that something is enclosed in a container. This need not be a well-defined three-dimensional enclosure with sharp boundaries. Nevertheless, MEs usually are temporally and spatially bounded—or contained—events. Where speakers express doing activities *in* a certain type of weather, they experience the weather event as the container. Contact is possibly also involved, such as the precipitate falling onto a FIGURE. In a similar manner, the preposition *durch* 'through' can put the speaker / FIGURE right in the middle of an ME; however, the focus is on the PATH that the FIGURE takes through the ME. Example (324) nicely demonstrates the usage of *im Regen* 'in the rain' and *durch den Regen* 'through the rain', where the latter may also appear more durative and prolonged. Ultimately, though, the main event predicate, and/or any temporal modifiers, determine the duration.

(324) *Und obwohl mich der Busfahrer gesehen hat, zog er es ebenfalls lieber vor, die Türen zu schließen, um seinen Kunden **im Regen** stehen zu lassen. Poah, war ich sauer! Aber was will man machen? Also lief ich **durch den Regen** weiter.* (elthas.de)
'And although the bus driver saw me, he preferred to close the doors to leave his customer out **in the rain**. Man, was I mad! But what can you do? So I kept walking **through the rain**.'

Example (325) similarly focuses on the PATH the FIGURE takes for "kilometres" without specifying an endpoint, a goal, or the speaker eventually exiting the ME, be it because of the rain stopping or entering a "different container", like a building or a rain-free area.

(325) *Weil meine Begleiter schon vor halb drei Schluss machen, … muss ich am nächsten Tag wohl oder übel **durch den Regen** laufen, einen Kilometer nach dem anderen.* (fairy-club.de)
'Because my companions finish before half past two …, I have to walk **through the rain** the next day, for better or for worse, one kilometre after the other.'

In summary, this section concerned a set of [PREP+ARG] constructions involving the *am*- and *beim*-progressive as well as constructions involving the prepositions *bei* 'at', which usually occurs with bare nouns, as well as the prepositions *im* 'in the' and to a lesser degree also *durch* 'through'. The constructions involving the progressives are considered too informal to be

licensed by language authorities like the *Duden*, but they are frequently used by speakers and newspapers, as corroborated by the findings in the *deTenTen13* and *DWDS* corpora. Since these constructions generally translate to 'when/if or while it rains/is raining', they appear to be valid substitutes or parallel constructions to the predicate construction. As German does not require aspect marking, the progressives make the utterances with the respective ME very immediate. Furthermore, they can profile the ME itself rather than being backgrounded, as is the case with the purely prepositional constructions with *bei* 'at' and *im* 'in the', which do not constitute independent clauses and usually provide the backdrop to more profiled activities or events. Nevertheless, as the equivalent of a German subordinate clause, they are a simpler, more felicitous linguistic solution than a more cumbersome subordinate clause. Moreover, an English translation that is closer in form to the German construction may lead to a literary or even poetic style (cf. Table 104 [i] *bei dem gestrigen Regen,* literally 'at/in yesterday's rain' compared to 'when it was raining yesterday'). This is not necessarily the case for the German construction, although the [PREP+ME_{ARG}] construction may be used in more formal or written contexts.

I reproduced the different uses of *bei* 'at' as per the *DWDS* definitions in Table 104 and found that the preposition is used frequently to express hypothetical or potential situations where the precipitates *Regen* 'rain', *Schnee* 'snow', and *Hagel* 'hail' are concerned. In contrast, the *Schneeregen* 'sleet' precipitate associates more often with expressions of simultaneity of two or more events and or activities, rather than conditionality. This can be explained by the fact that *Schneeregen* is not a major precipitate, but a mix of *Schnee* 'snow' and *Regen* 'rain' that is more difficult to forecast when a general rainy condition would subsume either precipitate type. There appears to be semantic overlap between the usage of the prepositions *bei* 'at' and *im* 'in the' where the latter appears to be construed as a container inside which the speaker and the weather event exist together, and possibly in contact with each other. *Bei* 'at', on the other hand, similarly happens simultaneously, but it does not necessarily affect the speaker or leaves them exposed to the ME. In like manner, the preposition *durch* 'through' resembles *im* 'in the' in its construal, but the focus lies on the PATH the speaker or the FIGURE takes in the container that is the ME.

The next section concerns idiosyncratic MEs, which frequently occur with the preposition *bei* 'at', but they may also profile an event. These idiosyncratic MEs are "bundles of MEs" wrapped up in two or three word compounds with *Wetter* 'weather', such as the previously mentioned *Bilderbuchwetter* 'picture book weather' in (315). These compounds exist

for overall "bad" and "good" weather conditions. The main focus of Section 4.3.3, however, lies in the eponymic compounds *Kaiserwetter* 'Emperor weather' and *Fritz-Walter-Wetter* 'Fritz-Walter-weather'.

4.3.3 Idiosyncratic MEs and -wetter *'weather' compounds*

This section concerns circumlocutions of weather conditions, quasi "all-inclusive" eponyms and compounds, which package either beautiful or bad weather into a single entity without having to resort to the usual constructional building blocks for MEs. Some of these items, *Fritz-Walter-Wetter* and *Kaiserwetter,* in particular, echo the agentive uses of *Frau Holle* and *Petrus* in that their reference to certain historical individuals is so strong that the ME is clear without having to specify the particular conditions like precipitation, temperature, and ambience that prevail. Other compounds consist of a metaphorical constituent in combination with *-wetter* 'weather' in order to describe the overall ME, some of which are vulgar, especially regarding bad weather conditions. Where the ME is a backgrounded event, which modifies some other event, it is often accompanied by the preposition *bei* 'at', as explored in the previous section. Where the ME is profiled, it may occur with all sorts of predicates, but one predicate that all *Wetter* 'weather' compounds appear to collocate with is *herrschen* 'to reign/rule' or 'to dominate/prevail'. This is true for the noun *Wetter* 'weather' itself, as well, which occurs as a bare noun 1,416,936 times in the *deTenTen13* corpus. As a subject, *Wetter* 'weather' collocates 1,521 times with *herrschen*, with a *logDice* of 6.7.[143] As an accusative object, *Wetter* 'weather' collocates 1,647 times with *herrschen*, and with a *logDice* of 7.3. Also, as an accusative object, the subject is likely to be the pronoun *es* 'it', as in (326). The ME is modified in several ways, but the crucial modification is that of quality (*sehr trüb* 'very dull/gloomy'), as *Wetter* 'weather' as a bare noun does not tend to *herrschen* 'reign' without specifying some kind of modification.

[143] The most frequent verb collocate for *Wetter* 'weather' as a subject is *mitspielen* 'permit' (lit. 'play along') with an incidence of 13,558 instances and a *logDice* of 11.2. (*Man versucht, soweit **das Wetter mitspielt**, so oft wie möglich an der frischen Luft zu sein.* 'You try to be out in the fresh air as often as possible, **weather permitting**.' It is also the most frequent verb collocate for *Wetter* 'weather' as an accusative object with 2,398 instances and a *logDice* of 9.2.

(326) ***Es herrscht weiterhin sehr trübes Wetter mit anhaltendem Regen****, der am Nachmittag in Schnee übergeht*. (DWDS, Münchner Merkur, 'Munich Mercurius' newspaper)
Lit. '**It continues to prevail very dull weather with persistent rain**, which, in the afternoon, changes into snow.'
'The weather continues to be very cloudy with persistent rain, turning into snow in the afternoon.' (Google translate)

It is also ungrammatical to rephrase the ME by using either the *am*- or *beim*-progressive (**das Wetter ist am herrschen* 'the weather is at the prevailing'). The least necessary modification is a determiner. In (327), a definite determiner and a relative clause modify *das Wetter* 'the weather'. The temporal modifier is not obligatory. Take note that the ME appears with the preposition *bei* 'at'. While *herrschen* appears to imply durativity, i.e., the event may possibly be going on for hours if not days, the modifier *zwischendurch* 'in between/from time to time' could imply shorter time periods. More data would have to be examined to gauge whether there is consensus between speakers regarding time frames or whether the type of weather is so dominant that *herrschen* 'rule/reign' appears like the default preferred predicate.

(327) *Jeden Tag hätte ich nicht zelten wollen,* ***bei dem Wetter****,* ***das*** *zwischendurch* ***geherrscht hat***. (outdoorseiten.de)
'I wouldn't have wanted to camp every day **at the**.DAT **weather that prevailed** from time to time.'

Since the weather compounds with *Wetter* 'weather' that are explored in the following sections are already modified by at least one of their compound parts, they do not require additional modification. The next section concerns "bad weather" compounds, where the focus is mostly on the eponymous *Fritz-Walter-Wetter* 'Fritz-Walter-weather'. Section 4.3.3.2 explores "good weather" compounds with a focus on *Kaiserwetter* 'Emperor weather'.

4.3.3.1 Fritz-Walter-Wetter *and other "bad weather" compounds*

Fritz-Walter-Wetter 'Fritz-Walter-weather' refers to the rainy conditions in which the German football legend, Friedrich "Fritz" Walter (1920–2002), played his best. Walter had contracted malaria during his time as a POW, which made him sensitive to heat, but his physical condition improved in lower temperatures and rainy weather. Most notably, Walter, as the team captain, led the German national team to victory in the 1954 World Cup, which was the first time that Germany was allowed to participate after WWII. In the finale, the German team ended up beating the heavily favoured Hungarian team against all odds. The match was played in heavy

rain, or *dem Fritz sei Wedder*, lit. 'of the Fritz his weather', as they said in the dialect of his native Kaiserslautern. Apparently, on the day of the finale, dark clouds gathered and it started to rain. Allegedly, Walter commented that "everything is clear now, nothing can go wrong" when his teammate Max Morlock cried out "*Friedrich, es regnet!*" 'Friedrich, it's raining!' (Jessen, Stahl, Eggers, and Schlüper 2003: 79). Winning the World Cup as the underdog provided a sorely needed boost to German morale and served as an inspiration to a divided nation. The victory is also known as "The Miracle of Bern". Undoubtedly, the significance and mythos of the event popularized and cemented the expression *Fritz-Walter-Wetter* or *dem Fritz sein Wetter* in the vernacular (a search on google.de for *Fritz-Walter-Wetter* returned more than 2.5 million results, the English equivalent *Fritz-Walter-Weather* still returned more than 1.5 million hits on google.com). Even if the expression may have lost some of its popularity in current everyday language, the seasoned football fan, at least, will be aware of it – and not just in Germany. The term has been translated into 26 languages on *Wikipedia*, it is part of the *FIFA* (*International Association Football Federation*) website, the *United Nations Climate Change* website,[144]as well as many other websites (in several languages) dedicated to football. Nevertheless, *Fritz-Walter-Wetter* does not have a *Duden* entry, but the *Dudenverlag* 'Duden publishing house' tweeted about the term and its meaning in drawing attention to the book *Die 50er – Ein Jahrzehnt in Worten* 'The 50s – A Decade in Words', which they published in 2019.[145] The term does have entries in the *DWDS* dictionary and *Wikipedia*, however, and the *DWDS* dictionary furthermore provides a word trajectory based on their newspaper corpus as given in Figure 44 with peaks of 64 items in 2014 and 65 items in 2004. Generally, the peaks occur in even years, which, coincidentally, is when either the Football World Cup or the *UEFA European Football Championships* take place (each every four years with two years between the events). Presumably, journalists employ the term *Fritz-Walter-Wetter* to a greater extent when reporting on football events with rainy weather conditions. The increase from 13 instances in 1998 to 32 instances in 2002 can possibly be explained by Fritz Walter's death in 2002, which perhaps "revived" the term and its usage regained some currency.

144 Source: unfccc.int/rain-in-bern-or-break-of-dawn-of-a-republic (last accessed April 20, 2024).
145 Source: twitter.com/Dudenverlag/status/1121670888695271424?lang=en (last accessed April 20, 2024).

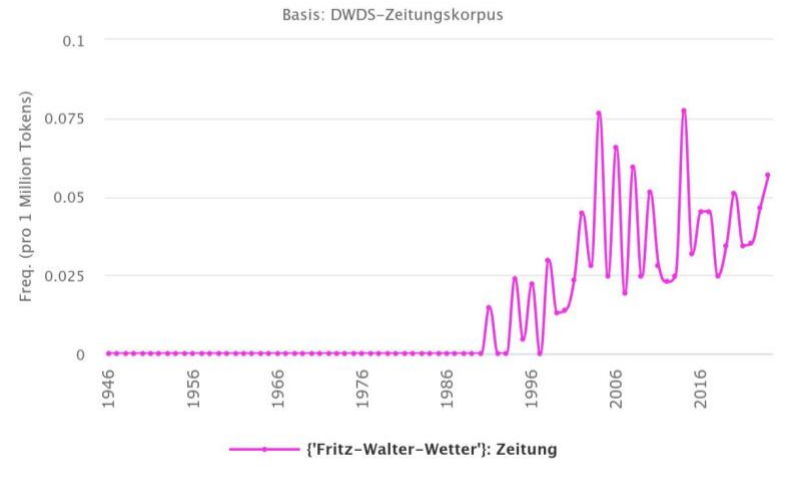

Figure 44. *Word trajectory (frequency per one million tokens) of the term* Fritz-Walter-Wetter *in the* DWDS *newspaper corpus from 1946–2016 (downloaded from the* DWDS*).*

The *DWDS Metakorpora* returned a total of 770 concordance lines for *Fritz-Walter-Wetter* where 107 are instances of *bei Fritz-Walter-Wetter* 'at Fritz-Walter-weather' with no modifiers between the preposition and the target term. The *deTenTen13* returns 339 concordance lines for *Fritz-Walter-Wetter* and another 72 concordance lines for the same term without hyphens. The prepositional construction *bei Fritz-Walter-Wetter* 'at Fritz-Walter-weather' occurs 73 times in the *deTenTen13*, with up to four modifiers allowed between the preposition and the hyphenated compound.

The majority of concordance lines, from both corpora, appear to be related to football events, either with reports on the game itself or football fans relating their experience as spectators, who have to endure the same weather conditions as the players. Frequent adjectival collocates are *bei bestem* 'at best', *feinstem* 'finest', *schönstem* 'beautiful', and *klassischem* 'classic'. The first three of these modifying phrases are interesting collocates with respect to rainy weather, as they are usually reserved for sunshine or blue skies. Examples with similar positive collocations are given in (328) and (329). The former is from a comment section on an upcoming football match in Wolfsburg and only consists of a rallying cry in addition to the remark on the weather. Despite *Fritz-Walter-Wetter* being associated with rain and what is generally perceived as "bad weather", the stance conveyed in the utterances does not appear to be negative, even where modifiers like *herrlich* 'splendid' or *bestes* 'best' are used, nor do they seem to be ironic. This is probably due to the positive connotation arising from Fritz Walter himself and the accomplishments he achieved.

(328) ***Herrliches <u>Fritz-Walter-Wetter</u>*** *heute!* (*DWDS*: comment section wolfs-blog.de)
'**Splendid <u>Fritz-Walter-weather</u>** today!'

The example in (329) is from a fan blog report, which does not contain any other comment on the weather either, i.e., no specification of the precipitation type or manner or the temperature. The *bei* 'at' preposition is employed, which backgrounds the ME while the trip to the stadium is profiled.

(329) ***<u>Bei</u> bestem <u>Fritz-Walter-Wetter</u>*** *ging es für uns früh ins Stadion*. (*DWDS*: 2hundert10.de)
'We went to the stadium early **<u>at the</u> best <u>Fritz-Walter-weather</u>**.'

As with *Frau Holle* and *Petrus*, there are examples that go into more detail describing the exact weather conditions, but the term *Fritz-Walter-Wetter*, too, is so salient and specific that additional details are not required or even obligatory. Example (330) contains the preposition *bei* 'at' as well. Furthermore, *Fritz-Walter-Wetter* is modified by the participle (*Partizip I*) *einsetzend* 'incipient', which, as an adjective, expresses simultaneity or an unfinished action. Take note that *Google Translate* returns a temporal phrase for the bolded expression ('when Fritz-Walter-weather set in').

(330) *Der Sturmlauf des Oberligisten blieb* ***<u>bei</u> einsetzendem <u>Fritz-Walter-Wetter</u>*** *aus.* (ftgern.de 'Sports Club Gern')
'The fourth division club's attack failed to happen **<u>at</u> incipient <u>Fritz-Walter-weather</u>**.'

In example (331), *Fritz-Walter-Wetter* is treated as an equivalent to *Regen* 'rain' as understood by the modifier *strömenden* 'pouring', since *Wetter* 'weather' itself cannot be modified by *strömend* 'pouring' (which sounds odd at best). By way of illustration, the *deTenTen13* returns only seven concordance lines for *strömend** Wetter* 'pouring weather', but 27,085 concordance lines for the collocation *strömend** Regen* 'pouring rain' where the asterisks do not reflect the actual search query and are merely placeholders for the different possible case inflections in German, rather than for full words between *strömend-* and a precipitate. Even a *Google* search only returns a total of five instances of *strömend** Wetter* 'pouring weather'.

(331) *So konnte die deutsche Mannschaft das Spiel offen halten und **bei strömenden Fritz-Walter-Wetter** erzielte Helmut Rahn knapp zehn Minuten später sogar den Ausgleich.* (blutgraetsche.de 'dangerous sliding tackle')
'The German team was able to keep the game open and **at pouring Fritz-Walter-weather**, Helmut Rahn even scored the equalizer just ten minutes later.'

While the examples above mostly employ the preposition *bei* 'at', the ME itself can also easily be profiled. In (332), this is done in combination with the verb *beginnen* 'begin', in a similar manner to (330), where, in contrast, the ME was backgrounded.

(332) *Pünktlich um 13:00 Uhr **begann** allerdings auch **das "Fritz Walter Wetter"** auf dem Achim Stöver Platz.* (tus-winterscheid.de 'Gymnastics and Sports Club Winterscheid')
'Nevertheless, punctually at 1:00 p.m. **the "Fritz-Walter-weather" began** on the Achim Stöver field.'

Since *Wetter* 'weather' is always going on, it would also appear odd to substitute *Fritz-Walter-Wetter* in (332) with a different type of *-wetter*, such as *Regenwetter* 'rainweather'. Here, too, the *deTenTen13* returns only four concordance lines, which contain the lemma *beginnen* ± four items + the word *Regenwetter*.[146] One of the four returns is given in (333).

(333) *In der Nacht **begann** dann wieder **Regenwetter***. (langstadt-aktuell.de 'Langstadt current')
'*In the night, **rainweather began** again*.'

A much more natural translation or, in fact, German equivalent for (333), would be constituted by either an argument (*begann der Regen wieder* 'began the rain again') or a predicate construction (*es begann wieder zu regnen* 'it started to rain again'). Again, this points to *Fritz-Walter-Wetter* being a substitute for the rain precipitate. Take note that rainy weather often indicates periods of rain, while *Fritz-Walter-Wetter* means rain outright. This inference of continuity is also made clear by the utterance in (334), which has to overtly specify the manner or frequency as *gelegentlich* 'occasional' with *Fritz-Walter-Wetter*.

(334) *Die Festivalbesucher konnten … nach Herzenslust kicken und taten das ohne irgendeine Pause: sowohl **bei gelegentlichem Fritz-Walter-Wetter** wie auch nach Einbruch der Dunkelheit unter Flutlicht*. (bolzen-online.de)
'The festival visitors could play football to their hearts' content … and did so without any break: both **at occasional Fritz-Walter-weather** and after dark under floodlights.'

[146] The exact CQL was [lemma="beginnen"][]{0,4}[word="Regenwetter"] within < s/>

In example (335), *Fritz-Walter-Wetter* is explicitly equated with *Regen* 'rain'.

(335) ***Fritz-Walter-Wetter, also Regen**, heißt hier einfach Tivoli-Wetter*.[147] (*DWDS*: reviersport.de)
'**Fritz-Walter-weather, i.e., rain**, is simply called Tivoli-weather here.'

There are several other compound words that refer to bad weather, but they are more generic and not based on eponyms. As opposed to *Fritz-Walter-Wetter*, which is clearly associated with the rain precipitate, these compounds may imply any type of precipitation. Table 106 shows a list of compounds associated with bad weather and their frequencies in the *deTenTen13* and the *DWDS Metakorpora* (the sum of the returns of each of the five metacorpora) as well as whether they have a *Duden* entry or not.

Table 106. *Frequencies of German "bad weather" compounds in the* deTenTen13 *and the* DWDS *corpora and their acceptance by* Duden online

German compound	literal translation	*Duden*	*deTenTen13* corpus	*DWDS Metakorpora*	*Total*
Schlechtwetter	'bad weather'	yes	8,709	47,198	**55,907**
Schmuddelwetter	'dirty weather'	yes	6,160	13,792	**19,952**
Sauwetter	'sow weather'	yes	4,411	6,370	**10,781**
Schietwetter	'shit weather' (regional)	no	1,873	4,413	**6,286**
Mistwetter	'crap/dung weather'	yes	2,554	3,331	**5,885**
Scheißwetter	'shit weather'	yes	672	952	**1,624**
Hundewetter	'dog weather'	yes	467	1,043	**1,510**
Schweinewetter	'pig weather'	no	95	109	**204**
Katastrophenwetter	'catastrophe weather'	no	57	88	**145**
Total			**22,444**	**73,965**	**96,409**

The majority of the compounds are colloquial, if not vulgar. It must be pointed out that the most generic and most frequent term, *Schlechtwetter* 'bad weather', as in (336), is still surpassed in frequency by the phrasal rather than lexical arrangement of its component parts, where the noun is modified by the adjective: *schlechtes Wetter* 'bad weather', as in (337). The *DWDS Metakorpora* alone return more than 250,000 concordance lines for the search query *schlechtes Wetter* 'bad weather'.

(336) *Nachdem auch noch **Schlechtwetter eingesetzt hatte**, mussten sie die Tour abbrechen und in das Basislager zurückkehren*. (*DWDS*,: *Wikipedia* entry on climber *Bernd Arnold*, 2004)
'After **bad weather had set in**, they had to break off the tour and return to the base camp.'

(337) *Zum Zeitpunkt des Unglücks **herrschte schlechtes Wetter***. (*DWDS*: *Wikipedia* entry on an Air Canada accident, 2024)
'At the time of the accident **bad weather prevailed**.'

[147] The Tivoli is a football stadium in Aachen, Germany.

The terms *Schmuddelwetter* 'dirty weather' and *Schietwetter* 'shit weather' are associated with the north and, according to the *DWDS Wortprofil* 'word profile', mainly collocate with the adjectives *norddeutsch* 'northern German' and *hamburger* 'of Hamburg'.[148] The *DWDS* dictionary literally uses an example from a northern German newspaper to describe the meaning of *Schietwetter* 'shit weather', as in (338). Take note that the utterance uses the *es+herrschen* construction.

(338) ***Es herrscht Schietwetter****, wie der Norddeutsche sagt: Regen, Wolken und starker Wind.*
(*DWDS* dictionary, *Der Prignitzer*, 29.05.2015, regional newspaper)
'**It prevails *Schietwetter***, as the northern German says: rain, clouds, and strong winds.'

The first noun of these compounds tends to refer to either bodily waste or an animal, i.e., *Schwein* 'pig' and *Sau* 'sow', often also associated with dirt or filth.[149] *Hundewetter* 'dog weather', as opposed to the other compounds, returned a fairly high number of instances in the *DWDS* metacorpus *Historical Corpora* (N=73), as in (339), which is taken from a 1890 story named *Seifenblasen* 'soap bubbles'. *Google Translate* chooses the translation 'terrible weather' for the term *Hundewetter* 'dog weather'. The second clause clarifies that the protagonist is referring to snowfall.

(339) *„**Ein Hundewetter**," sagte der Stadtrat. „Diesen Schnee wieder hinauszuschaffen, kostet der Stadt."*
"**A dog weather**," said the city councilor. "It will cost the city to get this snow out again."

The more vulgar term *Sauwetter* 'sow weather' also occurs in the historical corpora, where an utterance from a grade nine textbook specifies the types of weather further and where *Google* translates *Sauwetter* as 'a real bad weather'.

---

[148] *Schiet* 'shit' comes from the Low German for *Scheiße* 'shit'. Interestingly, *Schiet.*MASC and *Schiete.*FEM (both 'shit') have their own entry in the *Duden online* while *Schietwetter* or *Schietwedder* (both 'shit weather') do not.

[149] The nouns *Schwein* 'pig' or *Sau* 'sow' are frequently associated with dirt, filth, or untidiness and add emphasis to some German expressions. To name just a few, like in English where an untidy room can be a 'pig sty', it is a *Schweinestall* in German. Terrible handwriting is a *Sauklaue*, lit. 'sow's claw', something that looks '*wie Sau*' 'like sow' is dirty or untidy, when it is 'really cold', it can be referred to as *saukalt* 'sow cold' (N=5,447 in the *deTenTen13*) or *schweinekalt* lit. 'pig cold' (N=1,685 in the *deTenTen13*); equivalents exist for 'really hot', albeit these are much less frequent.

(340) *Bald gießt es in Strömen, kräftige Schnee- und Hagelböen setzen ein:* ***ein richtiges Sauwetter*** *also.* (*DWDS*: *Deutsches Lesebuch für höhere Lehranstalten* 'German reading book for higher education institutions', 1914)
'Soon it is pouring rain and heavy squalls of snow and hail set in: **a real sow weather**.'

Since bad weather compounds already encode the speaker's stance and the "quality" of the weather, MEs containing these compounds usually do not require any additional modification. Where there is additional modification, it often appears to be a single intensifier, such as *richtig* 'real' in (340).

Lastly, I explored compounds with the canonical precipitates, *Regen* 'rain', *Schnee* 'snow', *Schneeregen* 'sleet', and *Hagel* 'hail, that have been the focus of most of this dissertation. The frequencies of these items in different weather compounds as found in the *deTenTen13* and the *DWDS Metakorpora* (summed up) are found in Table 107. For each corpus, the first row shows the bare noun frequencies and the second row contains the frequencies returned for the noun preceded by the preposition *bei* 'at' with no modifiers between the preposition and the compound. Unsurprisingly, *Regenwetter* 'rain weather' has the highest incidence by a wide margin, while *Schneeregenwetter* 'sleet weather' is the least frequent. *Schneewetter* 'snow weather' and *Hagelwetter* 'hail weather' are about equally frequent. Roughly one third of the *Regenwetter* 'rain weather' instances in the *deTenTen13* corpus co-occur with preposition *bei* 'at' according to a simple CQL search that allows for no modifiers. The *Word Sketch* function even returns 9,023 instances of *bei* 'at' occuring with *Regenwetter* 'rain weather'. More than a quarter of the *DWDS Metakorpora* returns contain instances of *bei Regenwetter* 'at rain weather'.

Table 107. *Frequencies of different compounds of precipitate plus* -wetter *'weather' and the corresponding prepositional construction* bei …-wetter *'at …-weather' in the* deTenTen13 *corpus and the* DWDS Metakorpora

Corpus	***Regenwetter*** 'rain weather'	***Schneewetter*** 'snow weather'	***Hagelwetter*** 'hail weather'	***Schneeregenwetter*** 'sleet weather'	***regnerisch** Wetter*** 'rainy weather'
deTenTen13					
bare noun	21,053	523	439	15	6,727
with **bei** *'at'*	7,201	52	4	0	2,130
DWDS					
bare noun	49,405	1,224	1,604	32	17,745
with **bei** *'at'*	18,835	86	11	2	1,818

As opposed to *Schlechtwetter* 'bad weather', where the adjective noun combination *schlechtes Wetter* 'bad weather' was much more frequent than the compound itself, *Regenwetter* 'rain

weather' is much more frequent than its adjective+noun counterpart *regnerisches Wetter* 'rainy weather'.

The next section explores the opposite of *Fritz-Walter-Wetter*, i.e., dry, sunny weather with blue skies and warmer temperatures, with a focus on the ME eponyms *Kaiserwetter* 'Emperor weather' or *Königswetter* 'King's weather' and other more generic compounds.

4.3.3.2 Kaiserwetter *and other "Bluebird Day" expressions*

The word *Kaiserwetter* 'Emperor weather', also *Königswetter* 'King's weather', basically implies the opposite of *Fritz-Walter-Wetter* and means a sunny day with blue skies. There are two possible historical sources for *Kaiserwetter* from overlapping time periods and with related explanations: Austrian Emperor Franz Joseph I (1830–1916) and German Emperor Wilhelm II (1859–1941). Kaiser Franz Joseph's birthday (August 18) was said to always be sunny (Paul 1992; *Duden online*). Similarly, any national festivities or ceremonies at which Kaiser Wilhelm made an appearance apparently were accompanied by blue skies and sunshine as well, so that weather of this type has come to be known as *Kaiserwetter* (Wander 1987 [1870]). During the Third Reich, the expression's formulation and connotations were briefly transferred to Adolf Hitler and *Kaiserwetter* was replaced by *Hitlerwetter* or, respectively, *Führerwetter* (possibly by the Minister of Propaganda, Joseph Goebbels, himself). The latter two expressions have, fortunately, fallen out of use while *Kaiserwetter*, and to a lesser extent *Königswetter*, are still common. There are ten returns for *Hitlerwetter* in the *deTenTen13*, six of which refer to a 2017 book title ("*Und morgen gibt es Hitlerwetter*", lit. 'And tomorrow it gives Hitlerweather') and seven returns for *Führerwetter*. The *DWDS* metacorpora return a total of 25 concordance lines for *Hitlerwetter.* Here, too, the majority refer to the book title (N=16), while *Führerwetter* only amounts to four utterances. Frequencies for each of these terms were obtained from *deTenTen13* and *DWDS* metacorpora search queries and are given in Table 108. Since there are more generic compounds, I also searched for *Bilderbuchwetter* 'picture book weather' (cf. example (315)) and *Sonnenwetter* 'sunweather', the latter as an equivalent to *Regenwetter* 'rainweather', as well as for the adjectival equivalent *sonniges Wetter* 'sunny weather'. Since the bare noun itself does not necessarily indicate an actualized ME, and there is no imaginable limit for the use with a predicate, I also searched for collocations with the preposition *bei* 'at', which, as we have seen, backgrounds the ME. For each corpus, the first column shows the bare noun frequencies and the second column contains the frequencies returned for the noun preceded by

preposition *bei* 'at'. For the *deTenTen13* results, up to four modifiers may occur between the preposition and the target word (except for the last row, which additionally contained the adjective *sonnig* 'sunny').[150] The *DWDS* search queries allowed for no additional items between preposition *bei* 'at' and the target compound.

Table 108. *Frequencies of different compounds with* -wetter *'weather' and the corresponding prepositional construction* bei …-wetter *'at …-weather' in the* deTenTen13 *corpus and the* DWDS Metakorpora

***Wetter*-compound**	**Corpus**	***deTenTen13* *bare noun***	***deTenTen13* *bei*** 'at'	***DWDS* *bare noun***	***DWDS* *bei*** 'at'	**Total**
Kaiserwetter	'Emperor weather'	7,306	2,162	18,474	6,435	**34,377**
Königswetter	'King's weather'	193	53	412	96	**754**
Traumwetter	'dream weather'	4,530	977	8,747	2,178	**16,432**
Bilderbuchwetter	'picture book weather'	1,476	311	4,382	1,151	**7,320**
Bombenwetter	'bomb weather'	688	64	1,047	97	**1,896**
Postkartenwetter	'postcard weather'	250	44	344	95	**733**
Sonnenwetter	'sun weather'	2,775	875	3,775	313	**7,738**
sonnig** Wetter	'sunny weather'	17,185	5,540	28,538	9,793	**61,056**
Total		**34,403**	**10,026**	**65,719**	**20,158**	**130,306**

Ignoring the last row for the time being, *Kaiserwetter* 'Emperor weather' has by far the highest incidence of positive -*wetter* compounds for the bare noun condition as well as for its occurrence with *bei* 'at' in either corpus. Otherwise, the use of the adjectival modification *sonniges Wetter* 'sunny weather', with and without *bei* 'at', is the most prolific. In contrast to *Fritz-Walter-Wetter*, *Kaiserwetter* associates with all sorts of "genres" and is not primarily restricted to football contexts. With respect to 'good weather' concepts, Andrason (2019: §3.2.3) mentions the lexical item *pogoda* 'good weather' for Polish, which "typically implies sunshine and a pleasant temperature" (2019: 92) and is a homonym of *pogoda* 'weather'. Take note that German also has the concept of 'weather' plus a diminutive, *Wetterchen* 'weather.DIM', which generally refers to good weather (*Was für ein Wetterchen!* 'What a [great] weather.DIM!')

In (341) and (342), the ME is foregrounded by the use of the predicate *herrschen* 'rule/reign' or to 'prevail/dominate'. Undeniably, for both *Kaiserwetter* 'Emperor weather' and *Königswetter* 'King's weather' the predicate *herrschen* seems an especially appropriate

[150] The exact query was, for example: [word="bei"][]{0,4}[word="Kaiserwetter"] within < s/>, where the third pair of brackets contained the respective fair-weather compound. The search query for the last column was [word="bei"][]{0,4}[lemma="sonnig"][word="Wetter"].

collocate. *Google Translate* chose to translate the MEs with 'it was imperial weather' and 'the weather was great'.

(341) *Man kann es gar nicht anders sagen: in Wilhelmshaven **herrschte** heute **Kaiserwetter**.*
(friesischer-rundfunk.de 'Frisian broadcasting')
'There's no other way to put it: **Emperor weather ruled** in Wilhelmshaven today.'

(342) *An allen drei Tagen **herrschte Königswetter**, bei dem alle Spieler (…) beste Leistungen zeigen konnten.* (*DWDS*: *Südkurier* newspaper, dated 03.07.2002)
'**King's weather ruled** on all three days, in which all players (…) were able to show their best performances.'

Regarding the semantics of the other compounds, *Traumwetter* 'dream weather', *Bilderbuchwetter* 'picture book weather', and *Postkartenwetter* 'postcard weather' are similar in that they portray "unreal" or "dreamy" conditions. Postcards are thought to portray real views, but, obviously, these are also often overly idealized depictions of a place. *Bombenwetter* 'bomb weather' has a couple of associations. Bombs are potent and have an enormous impact. When something is "great" in English, it is "the bomb".[151] A fantastic mood or atmosphere in German is called *Bombenstimmung* 'bomb mood' and a great success is a *Bombenerfolg* 'bomb success'. As such, in the past, the interjections *Bombenwetter!* 'bomb weather' and *Donnerwetter!*, lit. 'thunder weather', apparently were used to express astonishment, surprise, anger, etc. *Donnerwetter!* is still used as an interjection to express surprise, anger, or appreciation. They loosely translate to 'Gosh!' or 'Golly!'. Allegedly, during war times, especially in WWII, the news did not use the term *Bombenwetter* 'bomb weather' – for valid reasons. More literally, *Bombenwetter*, i.e., good weather for dropping bombs meant clear skies, no clouds. This meaning is still retained in how the word is used today. The compound *Sonnenwetter* 'sun weather' was searched for as the counterpart to *Regenwetter* 'rain weather' and is less than ten times as frequent as the latter with respect to the bare noun and about 25 times less frequent where it occurs with *bei* 'at'. However, the adjectival modification *sonniges Wetter* 'sunny weather' is much more frequent than *regnerisches Wetter* 'rainy weather', which means that two different encoding strategies are employed for opposing weather conditions. For *Sonnenwetter* 'sun weather', too, one of the first predicate collocations to appear in the returns

[151] Obviously, something that "bombs" is bad in English and *Google Translate* apparently is not up to date on colloquialisms and chose to translate *Bombenwetter* (in context) as 'terrible weather'.

is *herrschen*, as in (343). Again, *Google Translate* chooses to translate the compound with adjectives.

(343) *Während des Laufes **herrschte** klares **Sonnenwetter**.* (*DWDS*, *Der Reporter* 'the reporter' regional weekly paper, dated May 23, 2019)
'During the run **prevailed** clear **sun** **weather**.'
'During the run the weather was clear and sunny.' (*Google Translate*)

Despite the frequent examples with the predicate *herrschen*, there are frequent collocations with the preposition *bei* 'at', as well, as can be seen in Table 108 where roughly one-fourth to one-third of most of the compounds collocate with *bei* 'at'.

In summary, as opposed to English, German has a wide variety of "prepackaged" weather compounds, which encode two contrasting types of weather: good and bad. The eponymous terms mostly focused on here, *Fritz-Walter-Wetter* and *Kaiserwetter*, are not fully analyzable without knowledge of the historical background or at least context of use. Generic terms like *Traumwetter* 'dream weather' or *Scheißwetter* 'shit weather', whose translation is common in English, too, are easily understood and require no explanation. Where *Google Translate* was tasked with a translation of the respective examples, it generally chose adjectives for the first part of the compound. Obviously, for the more straightforward compounds (*Regenwetter* 'rain weather', *Schneewetter* 'snow weather', *Sonnenwetter* 'sun weather, etc.) the adjective-plus-noun options exist, but they are not equally employed for the same type of weather with any frequency. The compounds from both ends of the spectrum frequently occur with the preposition *bei* 'at' as well as the predicate *herrschen* 'reign/rule/prevail', where *herrschen* possibly expresses durativity, but more research is needed to confirm this impression. Generally, the compounds do not seem to be accompanied by additional modifiers since their overall semantics are so specific, with respect to type of precipitate and other quantity, quality, or durativity connotations. Accordingly, some utterances contain extra information, which spells out the different weather conditions in order to subsume them then under a matching compound, but other utterances contain no additional information whatsoever and the compound as an interjection suffices as a comment on the current weather conditions. The next section continues with the topic of figurativity and MEs that has been introduced in this chapter. At the same time, it closes the circle and goes back to the beginnings as it concerns the figurative use of MEs in their canonical constructions, such as *es hagelte Kritik* 'it hailed criticism'.

4.4 Figurative uses of conventional MEs

Overall, this chapter concerns marked ME constructions, which align more closely with canonical clauses and phrase types in German. These ME constructions are considered marked because they employ at least two heavy elements in an intransitive clause or, in a transitive clause, they add a causal agent, who is responsible for the ME. In this section, I first return to some of those canonical ME expressions, which I examined comprehensively in Chapter 3, especially when they are being used figuratively. Additionally, I explore some objects that frequently collocate with metaphorical uses of *regnen* 'to rain' and *hageln* 'to hail' predicates. Neither *schneien* 'to snow' nor *schneeregen* 'sleet' seem to lend themselves much to being used figuratively. As shown in Chapter 3, *schneeregen* 'sleet' is not yet an established predicate and the argument noun seems not a canonical enough precipitate to feature in metaphor. The *schneien* 'to snow' predicate appears in only one specific form (*hereinschneien* 'snow in[to]') that is predominantly figurative, which is discussed below.

As a reminder, *Hagel* 'hail' was the foremost precipitate to be used figuratively and the only precipitate which produced figurative uses in all three canonical constructions. The frequencies of the figurative uses per precipitate type within the canonical constructions examined in Chapter 3 are given in Table 109. Each precipitate type is addressed separately below starting with the least frequent (*Schneeregen* 'sleet') to the most frequent (*Hagel* 'hail').

Table 109. *Frequencies of figurative use per precipitate and construction type based on the coded data frames downloaded from the* deTenTen13 *corpus examined in Chapter 3*

Precipitate	[*fallen* +ME_{ARG}]	[*es geben* + ME_{ARG}]	ME_{PRED}-Cx	Total
Regen 'rain'	13	0	8	21
Schnee 'snow'	0	0	4	4
Hagel 'hail'	10	6	175	191
Schneeregen 'sleet'	0	0	0	0
Total	**23**	**6**	**187**	**216**

The *Schneeregen* 'sleet' precipitate does not occur figuratively at all and, thus, is not discussed further in this section. The *Schnee* 'snow' precipitate does not seem to lend itself to figurative uses either. Both argument constructions returned zero figurative uses of the *Schnee* 'snow' precipitate. There are four instances of metaphorical use associated with the predicate construction based on the verb *hereinschneien* 'snow into'. *Hereinschneien* can be used literally

(*it snowed into the car*); nevertheless, the literal meaning appears to be infrequent. The figurative meaning that something or someone turned up unexpectedly is given in (344).

(344) ***Ihr könnt** einfach ohne Anmeldung **<u>hereinschneien</u>**.* (*DWDS*, klang-therapie-praxis.ch 'sound-therapy-practice')
'**You can** just **<u>snow in</u>** without registration.'

The first set of *hereinschneien* 'snow in(to)' results (N=50) in the *deTenTen13* does not contain a single result in which the meaning is literal. The four figurative examples in the data set that was downloaded and coded for the predicate construction are instances of 'something snowing into …' followed by objects like 'politics into everyday life', or 'an invoice into the house'. The oldest example of *hereinschneien* 'snow into' from the *DWDS* historical corpus is from 1830. It is figurative as well, so that non-literal uses have been around at least for 200 years or so. Presumably, snow is a more unexpected precipitate than rain, which may be the reason why *hereinschneien* 'snow into' rather than *hereinregnen* 'rain into' has come to mean 'something turning up unexpectedly'. Moreover, *Schnee* 'snow' is a precipitate that often has a positive connotation (cf. Section 3.2.2.4), which may be transferred to unexpected, but dear visitors. However, more data need to be examined to confirm this suspicion. Otherwise, *hereinschneien* 'snow into' may just as well be neutral in terms of its semantic prosody.

The *Regen* 'rain' precipitate produced no figurative returns for the [*es*+*geben*] construction. The [ARG+*fallen*] construction returned similar low frequencies for figurative uses of *Regen* 'rain' (N=13) as for *Hagel* 'hail' (N=10). Here, the precipitate mostly heads a modifying phrase as in an analogy (cf. ex. [114]: *fiel der Regen des globalen Kapitalismus* 'fell the rain of global capitalism'). The predicate construction returned eight instances of figurative uses of *regnen* 'to rain', where seven occurred in transitive constructions where "the thing which it rains" were instances of *Geld* 'money' (N=2), *Goldmedaillen* 'gold medals' (N=1), *Gold und allerlei Segnungen* 'gold and all kinds of blessings' (N=1), *Kartoffeln* 'potatoes' (N=1), but also *Kritik* 'criticism' (N=1) and *Zeitstrafen* 'time penalties' (N=1). The final concordance return was about an innkeeper who 'let it rain' from his watering can. From this rather small data set, it seems that the object of the verb *regnen* 'to rain' in transitive clauses is generally more positive and "plentiful". The last two objects, 'criticism' and 'time penalties' appear to be infrequent as a collocate of *Regen* 'rain' and tend to associate with the *Hagel* 'hail precipitate, as demonstrated below. The *Word Sketch* for *regnen* 'to rain' confirms that the more frequent accusative objects

of *regnen* 'to rain', or those with a stronger *logDice* score, are not necessarily negative prosodically. The most frequent collocate is *Bindfäden* 'threads/strings' (N= 1,098, *logDice*= 10.2), which simply means that it is 'pouring'. The second most frequent and second strongest collate is *Rosen* 'roses' (N= 913, *logDice*= 8.3). Many of these instances come from a German popular song called *Für mich soll's rote Rosen regnen* 'it shall rain red roses for me' by Hildegard Knef. The original from 1968 was rerecorded as a rock version in collaboration with the band Extrabreit and remained in the German charts for 22 weeks in 1993. Other strong collocates, that are not MEs, are *Hirn* 'brains' (N= 599, *logDice*= 7.9) and *Konfetti* 'confetti' (N= 186, *logDice*= 7.6). While the latter shares some similarities with actual rain in that it consists of many small particles, *Hirn* 'brains' is mostly part of a fixed expression, an invocation to the Lord to 'let brains rain from the sky' (*Herr, lass Hirn vom Himmel regnen*), when someone has said or done something particularly stupid. Apart from the nice triple alliteration, this expression also implies that rain is something good and fruitful. Speakers do not implore the Lord to "hit" people with brains, as would probably be the case for the *Hagel* 'hail' precipitate. Finally, there are the fairly frequent, if not necessarily strong, collocates *Katzen* 'cats' (N= 261, *logDice*= 4.9) and *Hunde* 'dogs' (N= 319, *logDice*= 3.5), which often appear in combination, since German allegedly borrowed the expression as a calque from the English 'raining cats and dogs'. That being said, German also has the expression *junge Hunde regnen* 'rain young dogs', as in (345), meaning 'to pour', which may explain the overall higher frequency of *Hunde* 'dogs'.

(345) *Wir wollen im Freien spielen, aber am Himmel hängen schwere graue Wolken, es nieselt greulich und der Wetterbericht kündigt an, daß* ***es junge Hunde und vielleicht sogar Schusterbuben regnen würde****.* (spiele.klausvogler.de 'games Klaus Vogler')
'We want to play outside, but heavy gray clouds are hanging in the sky, it is drizzling horribly and the weather forecast says that **it would rain young dogs and maybe even cobbler boys**.

The second object in this example, *Schusterbuben* or *Schusterjungen* 'cobbler apprentices' is less common, but still prevalent (N= 18, *logDice*= 4.3). The *DWDS* metacorpora returned 93 instances in total for the search expressions *Schusterbuben* && *regnen* and *Schusterjungen* && *regnen*. The expression is, obviously, less common and has different meanings depending on the region. The *Duden online* lists *Schusterjungen regnen* as slang from Berlin based on the assumption that there used to be many cobbler apprentices in Berlin. *Schusterjunge* is also the proper name for a (cheap) rye bread roll in Berlin, but why that should be related to precipitation is unclear, according to representative Manfred Glauninger from the *Institute for*

Lexicography of Austrian Dialects and Names.[152] In Austria, the expression *Schusterjungen regnen* apparently is purely metaphorical and means 'adverse circumstances' or 'to fight an uphill battle'. If you receive something *wenn's Schusterbuam regnet* 'when it is raining cobbler apprentices', it means that you will never receive it.[153] This, according to Glauninger, is connected to the figure of the quick-witted cobbler's apprentice, which apparently used to be very popular in older forms of Vienna theater. Apparently, the expression is much more common in the southern regions of Germany.

Apart from the utterances in the coded data set, as discussed in Chapter 3, there are many other metaphorical synonyms of *regnen* 'to rain' that will only be mentioned in passing here. The usage of these verbs often depends on the intensity of the rain as well as on the speaker's stance. Furthermore, many of these terms are not just colloquial, but outright vulgar. Some neutral verbs are *gießen* and *schütten*, both roughly translate as 'to pour'. The vulgar terms draw parallels to urinating as in *pissen* or *schiffen*. *Schiffen* goes back to the use of chamber pots that were shaped like a *Schiff* 'ship', where *schiffen* was first used for 'urinating' and then for falling rain. Moreover, *regnen* synonyms may depend on the region and their etymology may also be associated with what the rain looks or sounds like. A fine rain that looks like threads or strings is referred to as *fieseln* in the north and *fisseln* in the south, which is a callback to the already mentioned *Bindfäden regnen* 'it's raining threads/strings'. More onomatopoeic synonyms include *pladdern, plästern,* or *sicken*, which all seem to come from Low German words. *Pladdern* is a term used in the north of Germany, which goes back to the Low German noun *Kohpladder* 'cowpat' or 'cow droppings' or simply *Pladder*. Presumably, the sound that cow droppings make when they hit the ground is likened to big droplets hitting the ground. The verb *plästern* (Rhenish and Westphalian) similarly comes from Low German and means 'to spray with mortar' (cf. English 'to plaster' and Dutch *pleisteren*) where the sound component of the ME is likened to the sound of mortar being applied to a wall. *Sicken* may also mean to urinate, but in Rhenisch dialects it is apparently exclusively used for *regnen* 'to rain'. Again, according to the *Duden online*, it is a Low German variant of *seichen* vulg. 'urinate', which in turn comes from the Old High German *sihan* 'to flow softly trickling'. Thus, the onomatopoeic

---

[152] Source: www.derstandard.at/story/2110034/tiere-und-lehrlinge-fallen-vom-himmel (last accessed: May 7, 2024).

[153] *Bua* (SG) and *Buam* (PL) are Austrian variants of *Bub*(*e*) (SG) or *Buben* (PL), which translate to 'boy' or 'boys', respectively. The Standard German word is *Junge* (SG) or *Jungen* (PL). *Bub*/*Buben* is commonly used in Southern Germany.

synonyms for *regnen*, too, are largely related to excretion. Undoubtedly, there are many more synonyms for the *regnen* 'to rain' predicate.

Lastly, the *Hagel* 'hail' precipitate is the only precipitate type which produced figurative uses in all three canonical constructions. Moreover, the returns for the [*es+hageln*] construction were so plentiful that they made up nearly 88% of the coded concordance lines (cf. Section 3.4.3), and in every single figurative concordance line, the predicate was used transitively. Moreover, the majority of these utterances contained negative semantic prosody. The objects of the verb, i.e., "the thing which it hails", were varied. In *Word Sketch*, the strongest and most frequent collocation as the accusative object of *hageln* 'to hail' is the noun *Kritik* 'criticism' (*es hagelte Kritik* 'it hailed criticism'), with an incidence of N=4009 and a *logDice* score of 8.1. The remainder include accusative objects such as *Proteste* 'protests' (N=878, *logDice* 7.7) and *Absagen* 'cancellations' (N=545, *logDice* 7.7), *Zeitstrafen* 'time penalties' (N=111, *logDice* 7.3), or *Beschimpfungen* 'insults' (N=63, *logDice* 6.7). As already mentioned, both *Kritik* 'criticism' and *Zeitstrafen* 'time penalties' occurred as objects of *regnen* 'to rain' in the coded data where *hageln* 'to hail' appears to be the usual choice. It is possible that in these two examples, one ME was just inadvertently replaced for another. Furthermore, *Kritik* 'criticism' is also the most frequent grammatical subject occurring with verb *hageln* 'hail' (N=210, *logDice* 6.4) followed by *Schlag* 'hit/punch' (N=36, *logDice* 5.2). Nevertheless, there are also instances where the object of *hageln* 'to hail' has a positive (or neutral) connotation, such as *Komplimente* 'compliments' (N=83, *logDice* 5.4) or *Lob* 'praise' (N=119, *logDice* 4.9). Although the latter is more frequent than *Zeitstrafen* 'time penalties', its *logDice* is lower by 2.4 and, thus, the collocational association is considered to be weaker.

With respect to the figurative use of the argument constructions and the *Hagel* 'hail' precipitate, there does not seem to be much difference between the [*es geben Hagel*] construction, as in (346), and the [*Hagel fallen*] construction, as in (347).

(346) *Einige Zeit verstrich in der **es** nur noch **einen zarten Hagel an Schlägen gab***. (blogspot.de)
'Some time passed in which **it** only **gave a tender hail of punches**.'

(347) *Und so **fiel ein Hagel von Pfeilen** auf die Schutzschilder der Muslime*. (ansary.de)
'And so, **a hail of arrows fell** onto the protective shields of the Muslims.'

Moreover, the arguments are generally of same type as the most frequent collocates for the predicate *hageln* 'to hail', as given above: *ein Hagel von Zeitstrafen/ Schlägen/ Protesten/ Ermahnungen* 'a hail of time penalties/punches/protests/admonitions'. Take note that with

objects like *Schläge* 'punches' and *Pfeile* 'arrows', there are actual physical objects, which hit a GOAL. Some of the figurative utterances refer to actual projectiles falling from the sky like hail: *ein Hagel von Pfeilen*/*Steinen*/*eisernen Splittern* 'a hail of arrows/rocks/iron splinters'. Obviously, these uses indicate both that there are many of the respective projectiles coming down, but also that the temporality of the event is intermittent. These utterances contain analogies rather than metaphors. There is no actual physical impact of upset *letters to the editor*, *calls*, *criticism*, or *time penalties* when they arrive at their GOAL or to the RECIPIENT, at most they may hurt feelings. Other examples are *fällt der Hagel des göttlichen Zorns* 'falls the hail of divine wrath', which is also negative. The negative semantic prosody of the figurative examples of *Hagel* 'hail' thus goes hand-in-hand with the negative semantic prosody found in each of the literal concordance lines (bar a single concordance return) that contained a stance marker. Clearly, the majority of figurative examples exploits the salient, harmful quality associated with *Hagel* 'hail'. This makes more unusual collocates like *Komplimente* 'compliments' (N= 83, *logDice* 5.4) appear curious. Compliments are usually perceived as something positive, unless they are backhanded. With regard to the canonical constructions, the main takeaway is the overwhelming predominance of [*es*+*hageln*] instances that are used in a figurative manner, which constitute an unequivocal testament to the lack of equipotentiality among the precipitate types. This means, despite the sweeping typologically motivated generalisation that German uses the predicate construction for precipitation, it clearly does not do so equally for the four precipitate types. It is true, mostly, for the *Regen* 'rain' precipitate, but not for either the *Hagel* 'hail' precipitate or the *Schneeregen* 'sleet' precipitate. This outcome already revealed itself in Chapter 3 and a look at the potential of each precipitate to occur in a figurative context corroborates those findings. Obviously, the nature of each precipitate will influence how it is used figuratively, which cannot be elaborated on at this point due to space limits. However, since there seem to be few, if any, linguistic studies on weather metaphors in German, this may well be a worthwhile avenue for further research, with constructions with *Hagel* 'hail' especially.

4.5 Summary and conclusion

This chapter concerned marked ME constructions where the syntax is mostly unmarked, but the expression of the ME is not. This is achieved by using different strategies, which are

summarized in this section. None of the constructions in this chapter constitute a major encoding type, but their uses are nevertheless common and frequent in contemporary German. The choice of these constructions may depend on the genre, the precipitate type, or whether the speaker is profiling or backgrounding the respective ME. The encoding options for these marked constructions often forego the "light" element, like the pronoun *es* 'it', and, instead, employ at least two semantically heavy elements. Section 4.2 involved two construction types based on Eriksen et al. (2010), namely the cognate type and the split type. The cognate type was assumed to either be tautological, or to chiefly occur in marked genres like poetry. It was revealed, however, that the "simple" strategy of modification can make almost any cognate type ME felicitous, particularly in unmarked genres. Nevertheless, both the *Schneeregen* 'sleet' and the *Hagel* 'hail' precipitate are unlikely to appear in a cognate construction. For the former, this is due to the precipitate not just being the least canonical, but also the verb not having been accepted into the standard language yet. Additionally, the fact that *Schneeregen* 'sleet' is a multisyllabic compound may make it a less likely candidate for such a construction. While the *Regen* 'rain' and the *Schnee* 'snow' precipitates are used similarly, the latter, in particular, can easily appear as a resultative (*es hat 20cm Schnee geschneit* 'it snowed 20cm of snow') possible because the quantity of snowfall can be more easily gauged visually in comparison to *Regen* 'rain'.

The split construction also divides the semantic load between the predicate and the argument, but the two components focus on different facets of the ME. The ME itself, in German split constructions, is encoded in the argument. I focused chiefly on the use of predicate *rieseln* 'trickle/drift/flutter', which, nowadays, mainly occurs with the *Schnee* 'snow' precipitate. A close examination of the etymology divulged that it was originally associated with *Regen* 'rain', and I speculated that its current connection with *Schnee* 'snow' can be traced back to the popular German Christmas or winter song *Leise rieselt der Schnee*. The examination of *rieseln* lead to a brief exploration of its *anlaut* variant *nieseln* 'drizzle', which relates to fine *Regen* 'rain'. Both predicates can also occur as the first word of an ME compound, e.g., *Nieselregen* 'drizzle rain', which lead to a brief exploration of compounds with *Regen* 'rain', where the type of rain or the manner of the rain falling is encoded in the first part of the compound. These compounds were found to be part of professional terminology, but also part of vocabulary geared towards L1 and L2 language learners.

Section 4.3.1 examined the relationship of MEs with culturally relevant figures or artifacts in the guise of causal agentives, such as *Frau Holle* and *Petrus*. While these two figures are deities or saints in their own right, their association with specific precipitate types—*Frau Holle* with snow, *Petrus* with rain or absence of rain— is so salient that it requires invoking their names only to convey the weather situation somewhat accurately and without having to spell out the precipitate. For *Frau Holle*, this means 'shaking her beds', which translates to 'it is snowing'. Where she 'means well, there is plenty of snowfall. *Petrus*, on the other hand, as the gate keeper of heaven, is often associated with rainfall or dry, sunny days. Which one of the two it is, often depends on his "mindset": if he is merciless or does not mean well, it usually indicates rainy weather conditions. If is he understanding or sympathetic, or means well, it usually indicates the opposite of *Frau Holle* meaning well, i.e., zero precipitation; the weather is dry and sunny. These two culturally significant figures were explored in lieu of more generic atmospherical subjects, such as *der Himmel* 'the sky' or *die Wolken* 'the clouds'.

Section 4.3.2 concerned the prepositional use of MEs, which involved the "pseudo"-progressives *am*+precipitate and *beim*+precipitate as well as the "bare" prepositional use of (mostly) *bei* 'at' plus the precipitation type. The "pseudo"-progressives are in the process of becoming standardized. In a similar manner as the predicate construction, they may profile the ME, while the purely prepositional use with *bei* 'at' backgrounds the ME; at the same time another, concomitant event is profiled. All MEs introduced by *bei* 'at' imply simultaneity of two events, but, additionally, the relationship between them may be concessive, conditional, or causal. The data revealed that the *Hagel* 'hail' precipitate mostly occurred in conditionals, i.e., hypothetical conditions, which is not true for the other three precipitates. Furthermore, the perusal of the preposition *bei* 'at' revealed the use of other prepositions, such as *im* 'in the' and *durch* 'through'. Both appeared to impact the speaker/figure possibly physically more directly (e.g., getting wet), which is not necessarily given with *bei* 'at'.

Section 4.3.3 examined more weather compounds, but this time with *Wetter* 'weather' rather than *Regen* 'rain'. It tied in with the prepositional use, since the type of weather compounds, whether eponymous or generic, turned out to frequently appear with preposition *bei* 'at', such as *bei Kaiserwetter* 'at Emperor weather'. The compounds generally subsumed "good weather" situations and "bad weather" situations and I focused on two idiosyncratic, culturally relevant terms: *Fritz-Walter-Wetter*, which implies rainy weather, and *Kaiserwetter* 'Emperor weather', which implies sunny weather. Additionally, I briefly explored other, more

generic *Wetter* 'weather' compounds, whose first part generally employed a metaphor, which suggested the type of weather more transparently, such as *Schlechtwetter* 'bad weather'. The "bad weather" compounds also tended to employ bodily waste metaphors as the first part of the compound. The "good weather" compounds referred to idealized or "dreamy" places or states, such as *Traumwetter* 'dream weather'. Finally, this section also revealed the use of the occasionally frequently occurring predicate *herrschen* 'reign/rule' or 'prevail/dominate', which may also be used with the different precipitate types by themselves (i.e., not in a compound as in *zum Zeitpunkt des Absturzes herrschte heftiger Regen* 'at the time of the crash severe rain prevailed'). This makes *herrschen*, as a semantically heavy verb, a possible contender for the split type construction; however, its use with MEs is secondary.

Last but not least, Section 4.5 concerned figurative uses of MEs and returned to the canonical constructions examined comprehensively in Chapter 3 when used in a non-literal manner. A pronounced imbalance between the precipitate types was revealed, with the *Hagel* 'hail' precipitate appearing disproportionally often in metaphorical utterances in the predicate construction. Furthermore, all of these occurred in transitive clauses and contained almost exclusively negative semantic prosody with objects, such as 'criticism' and 'penalties'. The *Regen* 'rain' precipitate returned some of these more negatively construed objects, too, but also objects with a positive semantic prosody, such as 'money' and 'gold medals'. Overall, the predicate construction, as opposed to the argument constructions, returned concordance lines for three of the four precipitate types, i.e., *Hagel* 'hail', *Schnee* 'snow', and *Regen* 'rain'. The precipitate *Schneeregen* 'sleet' did not return any figurative results in any construction. The *Schnee* 'snow' precipitate occurred neither in argument constructions nor in transitive clause types and returned only one predicate construction being used metaphorically, which was *hereinschneien* 'snow in(to)'. *Hereinschneien*, used metaphorically, refers to someone or something showing up unexpectedly. Regarding the argument constructions, the *Regen* 'rain' and *Hagel* 'hail' precipitates appeared in similar contexts with negatively associated subjects/objects. However, the *Hagel+fallen* construction more often, within the small available sample size, appeared to contain actual physical objects, projectiles such as 'arrows' or 'splinters', that fall and not just metaphorical entities, such as 'criticism' or 'admonitions'.

Chapters 3 and 4 have provided a descriptive base of a wide range of German weather constructions, both unmarked and marked. The constructions examined in Chapter 4 are considered marked for different reasons, such as both major elements (the predicate and the

argument) being responsible for encoding some facet of the ME rather than the construction containing a single semantically light element or the "wholesale packaging" of weather conditions into a single lexical element. In Chapter 5, I provide a summary and conclusion of the dissertation at hand. I recapitulate the typological claims made about German (cf. Eriksen et al. 2010, 2012) and present the most pertinent results found in the data analysed in Chapters 3 and 4, how they relate to the typological claims and how they diverge with special consideration of the lexis-syntax interface and the field of lexical typology. Lastly, I provide a brief overview of how learners of 'German as foreign language' (*DaF*) are introduced to weather talk in light of findings of this dissertation.

CHAPTER 5

Whither the Weather?

This dissertation has presented a quantitative and qualitative analysis, based on corpus data, of meteorological event (ME) expressions in the German language with a particular focus on precipitation. The analyses in Chapter 3 were focused on three encoding types identified by Eriksen et al. (2010) in a major typological study of weather constructions: the argument type, the predicate type, and the argument-predicate type. Consequently, I examined three major constructions, which I termed "unmarked" for German (the ARG_{PREC}+*fallen* construction, the *es*+*geben*+ARG_{PREC} construction, and the *es*+V_{PREC} construction), in combination with four precipitate types (*Regen* 'rain', *Schnee* 'snow', *Schneeregen* 'sleet', and *Hagel* 'hail'). I also analysed several "marked" ME constructions, including the cognate type (*Regen regnet* 'rain rains'), the split type (*Schnee regnet* 'snow rains', *Schnee rieselt* 'snow flutters'), and other marked encoding types, such as causal agentives (embodied by *Frau Holle* 'Mother Hulda' and *Petrus* 'Saint Peter'), prepositional constructions (*bei Regen* 'at rain' and *am Regnen* 'at the raining'), as well as idiosyncratic MEs and *Wetter*- 'weather' compounds (*Kaiserwetter* 'emperor weather' and *Hundewetter* 'dog weather'). The data were chiefly sourced from the 16.5 billion-word web corpus *deTenTen13* plus the corpora available on the *DWDS*. The theoretical approach which served as the foundation for my analysis was cognitive linguistics and, more specifically, Langacker's Cognitive Grammar, as well as construction grammars in general.

My findings demonstrate that the choice of ME construction depends on contextual factors such as tense/aspect marking, temporal and locative modification, quality and quantity specification, figurativity, presence of stance-taking elements, genre and regional differences, as well as the individual precipitate type. For example, the prevalent German structure *es gibt* 'there is' (lit. 'it gives') is known to express absolute existence (*Es gibt einen Gott* 'There is a God'). However, in uses with MEs, it is predominantly found inflected in the past and future tense while less than one-fifth of the *es gibt Regen* 'there is rain' (lit. 'it gives rain') utterances express existential or generic propositions. At the same time, the PRECIPITATE+*fallen* ('PRECIPITATE +fall') argument-type construction seems to align chiefly with weather reporting, or as part of a technical meteorological genre, whereby information about the average weather

in a specific region is conveyed. This construction also turns up in texts on "how weather works", which goes hand-in-hand with its predominant use in existential or generic propositions, as well as in literary genres. However, the literary genre also attracts the argument-predicate type of ME construction, as in the cognate type (e.g. *Grauer Regen regnet dir aufs Dach* 'Grey rain is raining onto your roof') or split type (e.g. *Herbstspaziergang bei dem es Schnee regnete* 'Autumn walk where it rained snow').

Ultimately, there is no unitary distribution characterizing the interaction of precipitation type and construction type, a finding that echoes Andrason's (2019) and Andrason and Visser's (2019) findings for Polish and isiXhosa, respectively. The precipitate type *Hagel* 'hail', for example, overwhelmingly appears in the "default" predicate construction when the event is figurative and, furthermore, is largely imbued with negative semantic prosody, as in *es hagelte Kritik* 'it hailed criticism'. MEs expressing a literal use of hail associate most with the argument construction.

In this chapter, I will sketch out the implications of this study for typological research in general and the growing field of lexical typology in particular. I take the opportunity to review some weather participants, such as *Wind* 'wind' and *Nebel* 'fog' that do not involve precipitation and therefore were not otherwise addressed in this dissertation. Finally, I discuss what I believe are fruitful avenues for future research.

5.1 How we *actually* talk about the weather

The wealth of data examined in this dissertation proved to be a double-edged sword. For one, manually coding more than 100 corpus returns for each of 24 data frames (three constructions times two word orders times four precipitates) for 10 variables with up to six levels each was an ambitious project. At the same time, the corpus data proved to be a real treasure trove, which allowed for an extensive and detailed analysis of actual usage patterns, some of which would have stayed obscured had they not been revealed through the corpus searches. For example, neither the prepositional *am/bei*+ARG_{ME} constructions nor the [*es*+*gibt*+ARG_{ME}] constructions were identified by Eriksen et al. (2010), despite their prevalence in the language. Obviously, a typological approach that involves many languages does not allow for the depth of examination that was afforded to German precipitation events in this

dissertation, which is all the more reason to start with actual data—and lots of it. Using two different corpus hosts (*SE* and the *DWDS*) allowed me to explore more contemporary language use where many instances functioned as a proxy for spoken language (e.g., entries from message boards and private blogs in the *deTenTen13* corpus), while being able to countercheck many of these instances and usages against more formal weather language as provided by, for example, the newspaper corpora of the *DWDS*. Additionally, the *DWDS* historical corpus enabled me to investigate the development of items such as the split construction *Schnee+rieseln*, which appears to be a good example for how actual language use shapes linguistic knowledge and influences language change (Tomasello 2003).

What do we mean when we ask, "How do people actually talk about the weather?" Of course, the answer to the question about the verbalization or linguistic encoding of weather looks like everything else in language—it depends on context! To take Firth's (1957) dictum about the context-dependent nature of meaning further, it is not just the context of a word that matters, i.e., which words co-occur with it, but the extra-linguistic situation in which the weather talk takes place, including the genre and the speaker's attitude about the weather. Weather expressions do not exist in a vacuum. In fact, not incidentally, there is no weather in outer space, the closest physical approximation of a vacuum. If I look out the window while it is raining and I am asked, "What is the weather like right now?", would I say *es gibt Regen* 'it gives rain' or *es fällt Regen* 'rain is falling'? Probably neither of these responses would be appropriate. Following Eriksen et al.'s (2010, 2012) claim, I would probably resort to the PRED-Cx and say *es regnet* 'it's raining'. As the analysis of the three *Regen* 'rain' constructions has shown, the predicate construction not only is the construction most likely to associate with the present tense, it is also the most likely to occur with *Perfekt* tense, which is an indicator of oral speech, and it is the most likely to occur with a quasi-auxiliary that profiles information about when the rain event started or stopped. However, if was I asked to describe my current situation and environment more formally, I might say *es fällt Regen* 'rain is falling' or *draußen fällt Regen* 'outside falls rain' to specify that I, myself, am located inside a building or some other protective shelter. If I am with another person and we see that it's pouring rain, one of us might comment *Das regnet ganz schön* 'That is raining quite a bit' or *Da kommt ordentlich was runter!* 'It's really coming down' (lit. 'There comes really something down'), where the latter does not have to make mention of the rain precipitate at all. I am writing these words, coincidentally, *bei Regen* 'at rain', but, of course, the rain only provides the backdrop to the profiled event, i.e., the fact of

my writing. So, if I am asked, tomorrow, how the weather was yesterday, I might report *gestern gab es Regen* 'yesterday it gave rain' and *es soll noch bis Mittwoch regnerisch bleiben* 'and it is supposed to stay rainy until Wednesday'. This inventory already constitutes a wide variety of ME-constructions for what I would have coded as a WEATHER TALK situation.

Furthermore, consider the examples in (348)–(351), all taken from the *deTenTen13*, which contain mixtures of different constructions in combination with different precipitate types and other diurnal or temperature events. In (348), we find both a predicate construction and an argument construction along with two adjectival constructions. The examples in (349) and (350) feature both a predicate construction and an argument construction. In (351), the MEs are very elliptical and the speaker uses, in turn, a predicateless adjectival construction, a predicateless nominal construction with a quantifier (*viel Regen* 'lots of rain'), and a full predicate construction.

(348) *Ich laufe gerne **wenns regnet**, **dunkel ist**, **Schnee fällt**, **kalt ist**....also bei den vermeindlich* [sic] *schlechtesten Bedingungen.*
'I like to run **when it is raining**, **dark is**, **snow is falling**, **cold is** … so when it's supposedly the worst conditions.'

(349) *(…) das Arbeitsumfeld von ihr sieht immer (…) sauber aus. Selbst wenn **es regnet** oder **Schnee fällt**.*
'Her workplace/environment always looks clean. Even when **it is raining** or **snow is falling**.'

(350) ***Der Regen fiel** auf uns, **es hat geschneit**.*
'**The rain fell** on us, **it (has) snowed**.'

(351) ***Zu kalt**, **viel Regen** und vier Mal **schneite es**.*
'**Too cold**, **lots of rain** and **it snowed** four times.'

Obviously, speakers have no choice but to use a mix of constructions where parallel constructions are not available across all precipitates in a language. Nevertheless, where parallel constructions are available, speakers do not necessarily use the same type of construction throughout a list of MEs. More importantly, they do not default to the "default" predicate construction as demonstrated in the examples in (348)–(351) above. What is the explanation for using a mix of constructions rather than opting for parallel constructions that ostensibly represent the default? Diessel and Hilpert (2016) discuss priming as a significant factor in a speaker's choice where there are alternative constructions available. Priming is not just a matter of lexical access, but numerous investigations suggest it also affects "the speaker's selection of morphosyntactic structures" (ibid.: 14). Furthermore, a speaker's experience with

words and constructions influences their linguistic choices, which are fed by frequent co-occurrence patterns (ibid.: 16). Where there are alternative structures available, a speaker is more likely to repeat an already activated structure in the evolving discourse, which is possibly the case in example (31) in Section 1.5. It is impossible to know whether the authors of utterances which contain the same type of construction more than once (where several MEs are being listed) deliberately make stylistic choices or not, but there seem to be only two possibilities: (1) the speaker does not intentionally make stylistic choices and was either primed by a recent expression or simply opted for the most frequent or default construction; or (2) the speaker makes a conscious stylistic choice in order to make a salient comment. Arppe makes similar observations concerning these paralinguistic features, such as "attitude, emotion, and style" (2008: 243). In any case, we must assume that a speaker invokes a different construal with each choice of construction made. In the final analysis, "distributional knowledge is ultimately knowledge based on frequencies of occurrence, frequencies of co-occurrence, and dispersion characteristics", as Gries succinctly puts it (2012: 52). This includes modifying material such as the GOAL path in the argument construction in (350) and the quantifier in (351).

Eriksen et al. (2010: 587) make the following assumption: "the existence of the argument type is a trivial fact of languages; the existence of the predicate type (and the predicate-argument type) is not". Their reasoning is that if a language allows for nominals to encode meteorological participants, such as *snow(flakes)* or *rain(drops)*, it will be able to create argument type constructions, which can vary greatly in their productivity. They give the English examples *The rain is oozing from the clouds* and *Rain is approaching from the west* (ibid.) as novel argument ME constructions. This is a fair claim, which appears to be generally true. However, as soon as we consult authentic data we find, through frequency data, that there are strong preferences for certain collocations and interactions with specific extra-propositional and extra-linguistic variables. Knowing about these preferences, even—or especially—subconsciously, seems far from trivial. The analysed data revealed, for example, that the ARG-Cx with *fallen* 'fall', across precipitates, is the most likely to be associated with spatial modification, especially where the spatial modifier is a GOAL, a SOURCE, or a PATH (cf. example (350) above). At the same time, the ARG-Cx with *es geben* 'it give', is the most likely out of the three major constructions to occur with zero spatial modification, and where it does occur with spatial modification, it is mostly a LOCATION. Thus, we are not likely to encounter a sentence like *es gab Regen aus den Wolken* 'it gave rain from the clouds', which is not ungrammatical (cf. the

example (67) in Chapter 3), but certainly a bit odd. In fact, the most prototypical precipitation type, *Regen* 'rain', is the least likely to co-occur with a location (only about one-fourth of the returned instances contain a location). Similarly, the predicative *es+regnen* construction does not readily occur with spatial modification either unless it is a location; and when it is, it tends to be an adverbial deictic, such as *draußen* 'outside' or *hier*/*bei uns* 'here'. In contrast, the modifiers of place within the *Regen+fallen* construction data occur predominantly in a composite nominal phrase where deictics, too, usually co-occur with place-nouns (e.g., *hier im Süden* 'here in the south').

We could possibly conclude, then, from these co-occurrences that in using the PRED-Cx or the ARG-*geben* construction with the *Regen* 'rain' precipitate, the fact of rain is more important than exactly where it rained. This appears especially true for deictics like *draußen* 'outside', which is not very meaningful considering that *outside* is where weather events happen. Even *hier* 'here' or *bei uns* 'at our place/here' only tells the listener/reader that it rained where the speaker is without explicitly mentioning the place. In online discourse as on message boards and the like where users often stay anonymous, this information is basically akin to knowing that "somewhere" it is raining. Moreover, this type of extra-propositional information may be superfluous where interlocutors who know one another are interacting, as they will probably be aware of each other's location, but not of how long the weather event will last. This paucity of encoding locational information, of course, leaves open the opportunity for temporal modification, such as how long it has already been raining or when the rain started. Incidentally—and predictably—we are most likely to talk about MEs in the past. This is true for each of the three major constructions, but out of the three constructions, [*es*+*geben*+ARG_{PREC}] is the most likely to associate with the past tense. This finding was slightly surprising, as I originally suspected this construction to predominantly refer to future events. As it turns out, the [*es*+*geben*+ARG_{PREC}] construction occurs equally with either past or future tense and highly dis-prefers present tense.

In his analysis, Liégeois (2021) similarly shows how the linguistic properties of the text genre of weather reports (i.e., the information they convey and the corresponding terminology they select for), vary between target audiences. Interestingly, I found that the predicate construction is the least likely to associate with the genre WEATHER REPORT. This is an intriguing "paradox" with respect to Eriksen et al.'s (2010, 2012) claim: the genre type, which is literally dedicated to nothing but reporting on the weather is also the least likely to employ the

construction that assumedly is the most likely to be used when talking about the weather. These findings can only be reconciled if we appreciate not a tripartite typology, but a continuum that applies not just cross-linguistically, but also within a language. Hence, regarding the cross-linguistic implications of the present investigation, this study of German will hopefully serve as a model for others to carry out similarly in-depth, empirical investigations of weather expressions and adjacent domains in other languages based on authentic data. I anticipate that similar complexities and diversity will be revealed for other languages as well and subsequently be added to the growing field of the linguistics of the weather, which will refine Eriksen et al.'s (2010) tentative, and as of yet too coarse-grained typology. Clearly, as linguists we should strive to not only analyze authentic data, but to do so exhaustively beyond the bare bones of propositional content. In a truly cognitive fashion, a study of MEs should move away from overly structural approaches that chiefly focus on morphosyntactic features and consider, instead, critical factors such as speaker construal, frequency of use, context, genre, figurative conventions, and a host of extra-propositional collocational patterning that greatly affect choice of construction.

Not unlike the massive cross-linguistic research agenda spawned by Talmy's seminal work on motion events, Eriksen et al.'s (2010) typology may spark linguistic examinations of many languages, which will not just graze the surface, but provide profound, comprehensive assessments of MEs based on authentic, quantifiable data. This dissertation has made a modest start on German regarding the sub-domain of precipitation, but more research is needed with respect to other MEs. Originally, I intended to provide a comprehensive study of German weather events not limited to precipitation events, but including MEs such as wind, fog, atmospheric conditions, temperature, or diurnal expressions relating to dusk/dawn, etc. Such a wide-scale project proved too ambitious to be realized in a single dissertation. Nevertheless, the next section affords a glimpse into some of these heretofore neglected MEs.

5.2 Weather expressions in German beyond precipitation

Other than the expressions for precipitation, there are a host of weather expressions worthy of future research, some of which are explored in this section. In contrast to the examples such as *schneeregnen* 'to sleet', where verbs that are not "licensed" as such can

participate in the predicate construction, German has a couple of meteorological verbs (not denoting precipitation) that do not enter into the expletive predicate construction despite being listed as verbs in the *Duden*. Such verbs include *winden* 'to wind' and *nebeln* 'to fog'. To the author, *es windet* lit. 'it is winding' or *es nebelt* lit. 'it is fogging' seem awkward at best or at least highly marked as an ME option. In fact, one German speaker posted a question about the acceptability of *es nebelt* 'it is fogging' (with the disclaimer "German grammar") on the website *Gute Frage* 'good question', saying that her 3-year-old produced *es nebelt* 'it is fogging', which she found logical, but wanted to know why this was not a valid option in the German language.[154] Another user replied that her child's production was in fact valid or prescriptively correct, based on the fact that *nebeln* 'to fog' (and *winden* 'to wind') have their own entry in the *Duden*. Although these expletive predicate uses appear to be rare, examples can, of course be found, as the utterances in (352) and (353) for *nebeln* 'to fog' show.

(352) *Herr Pauli hat fast immer gute Laune morgens. Außer wenn **es** regnet oder **nebelt.***
'Mr. Pauli is almost always in a good mood in the mornings. Except for when **it is** raining or **fogging**.'

(353) *Sehen Sie aus dem Fenster und sagen "Egal, ob **es nebelt** oder regnet, es ist ein schöner Tag!*
'Look out the window and say "Doesn't matter whether **it is fogging** or raining, it is a nice day!'

A similar observation can be made for the meteorological meaning of *winden* 'to blow (wind)'. There are 91,654 results for the lemma *winden* on the *deTenTen13*, but only four out of the first 300 concordance lines refer to the meteorological event. Two examples are given in (354) and (355). The other 296 results translate to English 'wind/writhe/coil', as in streets, paths, or rivers that wind or to writhing in pain (*vor Schmerzen winden* 'writhe in pain').

(354) *Schließt die Augen und nehmt zunächst Euer Umfeld wahr. (...). Regnet es? Schwach oder heftig? **Windet es**? Schwach oder heftig?*
'Close your eyes and perceive your environment. Is it raining? Lightly or heavily? **Is the wind blowing** (lit. **is it winding**)? Lightly or heavily?'

(355) *Mittags um Zwei erreichen wir das Barrafu Camp. Barrafu bedeutet Schnee. Den hat es zwar nicht, aber **es windet** sehr stark, und es ist saukalt.*
'We reach the Barrafu Camp at midday 2pm. Barrafu means snow. That (snow) it does not have, but **the wind is blowing** (lit. **it is winding**) very strongly and it is freezing cold.'

[154] www.gutefrage.net/frage/warum-regnet-es-aber-nebelt-nicht (last accessed June 9, 2024).

Generally, though, the preferred option for encoding wind and fog in German seems to be the adjectival predicate type construction, as in *es ist ADJ* 'it is ADJ', exemplified in (356) and (357).

(356) *es ist nebelig* 'it is foggy'
(357) *es ist windig* 'it is windy'

The adjectival predicate type has already been briefly addressed in Section 1.5 where I singled out *schneeig/verschneit* – 'snowy' as opposed to *es ist regnerisch* 'it is rainy' (cf. also fn. 33). The adjective *schneeig* 'snowy' is not the only one that does not seem to be an option for the adjectival predicate construction, at least not with a meaning that parallels the semantics of 'it is rainy'. Though it has been claimed that *schneeig* 'snowy' is not used in German, counterexamples can be found, as in (358) and (359):

(358) *Der Frühling ist hier ziemlich kühl und der Sommer ist nass, warm (die aufgehende Sonne wärmt den Boden gut), regnerischer Herbst, der Winter ist nicht zu kalt,* ***es ist schneeig****, mit Tauwetter*.[155]
'Spring is fairly cool and summer is wet, warm (the rising sun warms the ground pretty well), rainy autumn, winter is not too cold, **it is snowy**, with thaw-weather.'

(359) *Hatte auch einige ganz prachtvoll schöne Tage, die ich zu Aufstiegen bis auf 8 bis 9000 Fuß benützte. Meist* ***war es*** *jedoch kalt und regnerisch und auch* ***schneeig****.* (uni-graz.at 'University Graz')
'(I) also had some wonderfully beautiful days, which I used to climb up to 8 to 9,000 feet. But mostly **it was** cold and rainy and also **snowy**.'

The latter example is taken from a letter written by Austrian physiologist and histologist, Alexander Rollett (1834–1903), to his wife. Since the MEs here are contained in a list (*kalt, regnerisch, schneeig* 'cold, rainy, snowy'), it appears that *schneeig* may not just mean 'snow covered', since a 'beautiful day' is not irreconcilable with snow-covered ground. Nevertheless, most utterances appear to confirm the *Duden* definition of *schneeig* 'snowy' meaning 'snow covered' rather than 'it is snowing (on and off)'.

In addition to *Schnee* 'snow', the precipitates *Hagel* 'hail' and *Schneeregen* 'sleet' do not participate in the adjectival predicate construction either: **es ist hagelig/hageli(s)ch* *'it is haily'

[155] Source: https://de.fehrplay.com/obrazovanie/83640-prirodnaya-zona-ukrainy-step-lesostep-smeshannye-lesa-gory.html (last accessed May 30, 2019). The *deTenTen13* returns 720 concordance lines for *schneeig* 'snowy' where it appears to mean snow covered, collocating with *Gipfel* 'summits', *Felder* 'fields', and *Boden* 'ground', but also the colour 'white'.

or **es ist schneeregnerisch* *'it is sleety', where the latter seems marginally acceptable to me. This assessment is presumably based on the fact that *regnerisch* 'rainy' is actually a derived word and *Schneeregen* 'snowrain' is a noun-noun compound. However, it seems that a day that could be called *schneeregnerisch* would have to fulfill certain conceptual prerequisites. Any day that could be perceived of as *regnerisch* 'rainy' would qualify; however, the temperature would need to be "at the brink of snow" – cold enough to not be rainy but warm enough to not be snowy. This seems to be a bit of a meteorological balancing act with a low strike rate for a speaker, unless the weather has been consistently *schneeregnerisch* 'sleety', possibly for several days in a row. Such an unlikely condition does not appear to warrant lexicalization. The analyses in Chapters 3 and 4 as well as these new examples have already demonstrated the way in which the German language carves up the weather lexicon. The next section elaborates on how the study at hand relates further to the field of lexical typology.

5.3 Revisiting lexical typology and implications for *DaF* 'German as a foreign language'

In this dissertation, I have addressed both the *onomasiological* perspective (i.e., which concepts require lexicalization or "world-to-the-word" mapping) and the *semasiological* (which concepts the word refers to or "word-to-the-world" mapping). As a speaker of German, the four main precipitate types, *Regen* 'rain', *Schnee* 'snow', *Hagel* 'hail', and *Schneeregen* 'sleet', seem to represent the four cardinal precipitation categories, despite the fact that *Schneeregen* 'sleet' is a (rather transparent) compound.[156] This is in contrast to *Graupel* 'graupel' which denote a softer form of hail that has little lexical or constructional currency. I considered whether *Schneeregen* 'sleet' may be a relatively novel compound and thus reduced in its productivity, but the *DWDS* word trajectory gives 1810 as the earliest token in their corpora. By contrast, the first token of *Graupel* 'graupel' stems from 1750. While the predicate *schneeregnen* 'to sleet' is not yet acceptable in standard German, the first token of the predicate *graupeln* 'to graupel', which is sanctioned by the *Duden*, appeared in a newspaper corpus from 1950. It should be noted that *Graupel* and *graupeln* were not considered in the present analysis. Perhaps the

[156] Take note that one of my examiners, Andrea Wilhelm, pointed out that *Graupel* 'graupel' is more cardinal to her than *Schneeregen* 'sleet'. This is possibly a regional difference between northern and southern Germany, which is not further addressed here.

presence in German of the predicate form *graupeln* 'to graupel' is a factor preventing *schneeregen* 'to sleet' from becoming a valid, standard option of encoding a sleet event with a predicate construction. Although, *graupeln* 'to graupel' occurs with an even lower incidence in the *deTenTen13* (N=85) than *schneeregnen* 'to sleet' (N=88). Perhaps it is simply the cognitive load of a longish compound predicate (cf. *es hat geschneeregnet* 'it has sleeted') that keeps *schneeregnen* 'to sleet' from becoming a frequently used, and thus standardized, item. Of course, how the ME lexicon is divvied up in different languages is fascinating. English has predicates for *hail*, *sleet*, and *graupel*, but speakers do not seem to employ them frequently. In fact, the item *graupel* appears to be rare if not fairly unknown in general. This assessment is based on personal communication. *Graupel*, sometimes also spelled *groppel*, is at least part of *The American Heritage Dictionary of the English Language* (5th edn.), *The Century Dictionary*, *Merriam Webster*, and *Collins Dictionary* (all online), as well as *Wiktionary*. Andrason and Visser (2019: 24), too, point out the heterogeneous distribution of verbal and nominal constructions for different precipitate types in isiXhosa: "there is no verb that could specifically and exclusively denote snow, hail, or sleet", which they attribute to extra-linguistic factors, such as the climate and the resulting differences in construal. Furthermore, as already noted, Visser (2019: 95–96) finds that Polish does not encode sleet as a singular lexical item, but rather that speakers must resort to a paraphrase of either 'rain with snow' (*deszcz ze śniegiem*) or 'snow with rain' (*śnieg z deszczem*).

In a similar manner to *Schneeregen* 'snowrain', I showed how German has all kinds of compounds which denote different types of rain (cf. Section 4.2.2.2), all of which conflate different characteristics such as quality (e.g., *Sprühregen* 'spray rain'), duration (*Landregen* [lit. 'land rain'] or *Dauerregen*, both meaning 'constant rain'), or intensity (*Platzregen* 'downpour'). Considering the German language's proclivity for compounding and the fact that *Graupel* is a "soft kind of hail", there do not seem to be restrictions which would prohibit a compound like *Weichhagel* 'softhail' or *Harthagel* 'hardhail'. It seems unlikely, then, that languages which already have to resort to periphrasis to talk about *sleet* would make space in their lexicon for a precipitate type like *Graupel*. It seems logical that the environmental realities of a speech community affect the lexicon. Where snow or sleet are not common or important factors of everyday life, there seems to be no necessity to encode these in lexical items but rather resort to circumlocution or paraphrase.

In the final analysis, Eriksen et al. (2010: 593) hypothesize that the scalar typology of p-encoding (cf. Figure 1) is also a visualization for the diachronic development of p-encoding from an argument focus to a predicate focus. This assumes that argument p-encoding—purportedly the most trivial and default way of describing an ME—develops over time into the most "customized for meteorological events" version on the right, which is predicate p-encoding. Andrason claims that his findings for Polish corroborate this hypothesis (2019: 100). It may also well hold for German. The noun *Graupel* 'graupel' (15th century), for example, is older than the verb *graupeln* 'to graupel' (17th century), according to the etymological dictionary of the *DWDS*.[157] The more restricted behaviour of the ME predicate *schneeregnen* 'to sleet' seems to be reflecting a delayed development, despite the first attested appearance of the compound *Schneeregen* 'sleet' over 200 years ago. However, as stated earlier, the sheer presence of *graupeln* 'to graupel' in the German lexicon may mitigate against the widespread adoption of *schneeregnen* 'to sleet' in predicate p-encoding constructions.

These diverging weather inventories obviously also have implications for L2 speakers and learners of the German language. In looking at *DaF* (*Deutsch als Fremdsprache* 'German as foreign language') materials online, it is not surprising that the learner worksheets for beginners exclusively present the possibilities given by Eriksen et al. (2010), i.e., the expletive predicate type, when it comes to learning how to talk about the weather. The worksheet *Wie ist das Wetter heute* 'how is the weather today' in Figure 45 lists the most prevalent weather participants (*Sonne* 'sun', *Wolke* 'cloud', *Regen* 'rain', *Gewitter* 'thunderstorm', *Wind* 'wind', and *Schnee* 'snow') and also presents the corresponding predicate constructions (*es regnet* 'it is raining', *es ist sonnig* 'it is sunny', etc.), which the learner can match to the respective pictures. These are probably the most basic options, or also the more prototypical MEs as neither *Hagel* 'hail', nor *Schneeregen* 'sleet', nor any of the argument constructions and items discussed in Chapter 4 are provided. By only reaching as deep as the canonical examples, the beginner material mirrors the descriptions made by typologists: it stays on the surface and allows for the

[157] *Graupel* also denotes 'grains of pearl barley', which is apparently the older meaning and possibly related to the meaning of 'soft hail'. Grimm's dictionary furthermore returns an intriguing historical example (from the German chronicles of Bohemia) for *Graupel* 'graupel' dated to the second half of the 16th century: *(den 22 tag jali) kam ein wetter, das es graupeln warf zu Trautnaw*, *als tauben eier* '(on the 22nd of July) came a weather that it threw graupel in Trautnaw, as pigeons' eggs'. Here, we have an agentive 'a weather', which 'throws' graupel. Moreover, we find a similar size comparison ('pigeons' eggs') as in the contemporary examples given in Chapter 3.

most basic expressions, but it has little to do with the natural expressions that fluent speakers use.

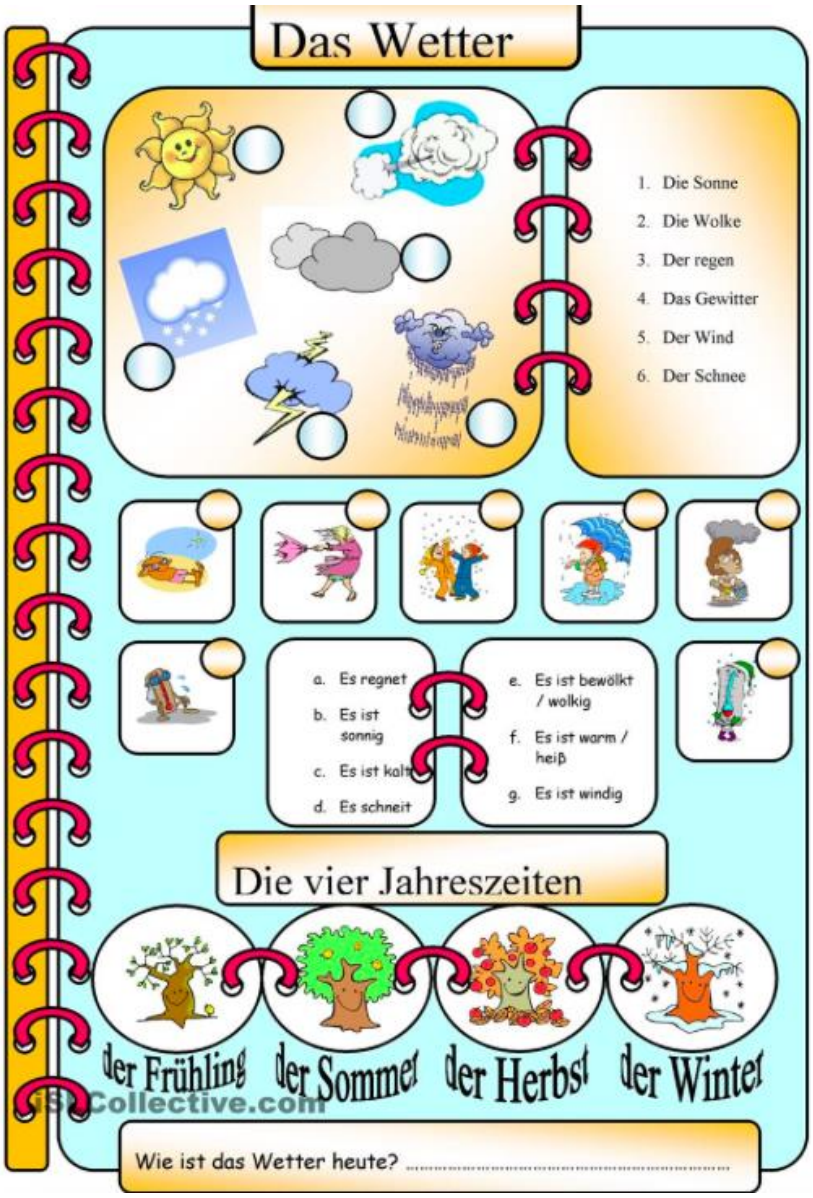

***Figure 45**. Exercise sheet lexical field "The weather" for beginning learners of German*[158]

[158] Source: https://de.islcollective.com (last accessed June 12, 2024)

Nevertheless, in materials for more advanced learners, options like the [*es gibt*+ARG_{PREC}] construction are included (cf. Figure 46). This is also true for reference/pedagogical grammars of German (e.g., Durrell 2011). The left side of the intermediate learner sheet titled *Wetterbesprechung* 'weather talk', given in Figure 46, contains the more formal, schematic items to talk about weather. The "column" on the right, however, gives actual examples that a fluent speaker of German might utter in a conversation about the weather. While the items on the left better resemble the items given in Eriksen et al.'s (2010) typology, the items on the right appear much more grounded in actual language use. Consider *Heute gibt es sicherlich noch Regen* 'It is surely (still) giving rain today', which expresses the speaker's conviction that it is going to rain later. It is formed with the [*es*+*geben*] construction, which implies an imminent ME and also contains an expression of stance. Additionally, the sheet exemplifies *Wetter* 'weather' compounds like *Superwetter* 'super weather', *Bilderbuchwetter* 'picture book weather', *Mistwetter* 'crap weather', and *Sauwetter* 'sow weather', which I explored in Section 4.3.3. Take note that these items are not necessarily embedded in a full sentence, but are merely exclamations, which are sometimes modified (*So ein…!* 'Such a ….!). This extends also to the non-precipitate items like *Ziemlich grau heute!* 'Pretty grey today!'. And here, too, we find that the items *nebelig* (here given as *neblich*) 'foggy', *windig* 'windy', and *stürmisch* are presented in the adjectival predicate construction although both items enter the expletive predicate construction as sanctioned by the *Duden* (*es nebelt* 'it is fogging', *es windet* 'it is winding', *es stürmt* 'it storms', cf. also Section 5.2 above).

In summary, for precipitation, the primary ME construction taught to beginners is predicate p-encoding with an expletive as the subject, at least for the more prototypical precipitation types, as in *es regnet* 'it is raining' and *es schneit* 'it is snowing'. Obviously, the FIGURE or "stuff" itself, *Regen* 'rain' or *Schnee* 'snow', is also taught, but it is not necessarily embedded in an actual example sentence. Moreover, any modifying material is largely ignored. The material for intermediate learners included much more varied and native-like options to talk about the weather including statements with an empathic reading, such as expressing either surprise or strong positive or negative feelings about the weather. These advanced materials more closely resemble the authentic data analysed and discussed in Chapters 3 and 4 and better showcase for learners the heterogeneity of encoding options regarding weather talk. Since these are options offered to intermediate learners—not even advanced—they should certainly not just be exemplified but also seriously considered in a comprehensive study of the

domain of weather expressions, be it typological, pedagogical, or descriptive—as I have attempted here.

Figure 46. *Exercise sheet lexical field "The weather" for intermediate learners of German*[159]

[159] Source: https://de.islcollective.com (last accessed June 12, 2024)

5.4 Limitations of typologies and the present study

The main insufficiency with Eriksen et al.'s (2010, 2012) typology, or rather with most typologies, lies in its assessment of a coarse-grained and shallow sample of a broad set of languages that is based on standardized options that do not always reflect actual language usage. The data that typically are considered in a typological study are often taken from secondary sources such as reference grammars and dictionaries. Personal knowledge of a language does not suffice either, since, and as stated in the first chapter, even educated speakers are not necessarily aware of the existence of certain items and structures, let alone their frequency or *preferred collocations.* Reference grammars, especially, tend to give only a few canonical or prototypical examples from the standard variety of a language. These data sources do not illustrate more peripheral options that native speakers would nevertheless have in their repertoire and use. Moreover, rarely are a wide range of lexical fillers for a given construction type ever considered. Accordingly, the authority on the standard use of the German language, the *Duden*, does not list the cognate type construction *Regen regnet* 'rain rains'; however, as noted in Chapter 4, the *DWDS* dictionary does. This is likely due to the fact that the *DWDS* dictionary is based on corpus data. Ideally, different methods and data sources are employed to collect, countercheck, and corroborate data (cf. Arppe and Järvikivi 2007). Because the nature of any typological study is to examine a robust and independent sample of languages, it follows that any resulting conclusions have to be taken with a grain of salt, as they are based on data that may be superficially drawn, highly normalized, and incomplete. Too often, typological studies are only concerned with the differences between the standard grammars of languages and central cases, thus disregarding actual usage patterns at the periphery. It is the periphery that best offers insights about actual language use. It is the periphery that best sheds light on the richness of possibilities that speakers of a language have at their disposal. And it is with actual usage data that we can understand how speakers make use of or even conventionalize the lexical items and constructions afforded by the language.

Traditional typological approaches can appear like pixelated images or photo mosaics made up of a vast number of smaller pictures. From a distance, what the observer perceives is a single image ("the big picture"), but closer examination reveals that that the big picture is, in fact, made up of hundreds or even thousands of smaller images of great variety. Lexical-

typological research must take care to avoid the shortcomings of traditional typologies, especially when it is the larger data sets that form the basis for lasting research. And here, too, the data are often not contextualized. This is precisely why a corpus-based approach is critical. Consequently, the present investigation provides evidence that Erikson et al.'s (2010) typology is not granular enough with regard to German, and therefore needs refinement.

While any analysis based on corpus data is fundamentally more veridical than one based on constructed data, corpora should be approached with caution as all are not created according to the same standards or intended to be used for the same purposes. As is true for each data set one analyzes in any study, a corpus is only fully representative of itself. Size, sampling procedures, amount of annotation, and ease of use are among those factors that researchers primarily take into account when selecting a corpus. It is always hoped that the selected corpus is at least somewhat representative of the speech community, whether or not it has fully sampled across speakers (ages/genders/dialects/socio-economic standings), modalities (spoken/written), registers (formal/informal), and genres (too numerous to mention). Perhaps the information I would have wanted most is on speaker demographics, such as age, sex/gender and level of education, but also which region or country the speakers are from and/or which dialect they speak (if applicable), since there is a plethora of different dialects in Germany alone (not counting Austria or Switzerland).

As already noted, the abundance of data led me to abandon some originally planned analyses, such as the examination of the hypernym *Niederschlag* 'precipitation'. Nevertheless, I downloaded the relevant data and made a first-pass annotation of the corpus returns for both argument constructions. The overall frequencies point to a preponderance of [*Niederschlag+fallen*] returns (N=3,921) over [*Niederschlag+geben*] returns (N=2,085). For the hypernym, we find a paucity of the predicate construction. The expression *es schlägt (sich) nieder* 'it precipitates (itself)' points toward a chemical process (as seems to be the case in English) rather than a meteorological event. Presumably this highly schematic and formal lexical item predominantly occurs in weather forecasting and meteorological texts and less so, if at all, in other genre types. Correspondingly, a cursory inspection of the data suggests that these utterances contain only minor expressions of stance, as one would expect from more formal genres. Thus, I assume that any mention of *Niederschlag* 'precipitation' would be followed up with the more specific precipitate type or types (depending on the region and altitude, that is to say, modifications of place), which would involve co-occurring meteorological events (or co-

MEs). Regarding the co-ME variable, I provided a provisional analysis of the co-MEs of the *Regen* 'rain' precipitate and the *Schneeregen* 'sleet' precipitate. For the latter, these often appear with irrealis markers, which indicates the probability of a forecast rather than the certainty of reporting on a past event. The *Hagel* 'hail' precipitate frequently occurs as co-ME, which is in all likelihood based on its natural occurrence as a "byproduct" of other MEs, such as thunderstorms and/or rain (cf. Section 3.4.1.5). For the *Regen* 'rain' precipitate, there is a stark difference with the occurrence of co-MEs in the argument constructions. In the [*Regen*+*fallen*] construction, they predominantly appear preceding the target ME and, to a lesser extent, following it. In the [*Regen*+*geben*] construction, they tend to occur as lists of multiple MEs. This is probably due to the fact that *geben* 'give' is semantically lighter than *fallen* 'fall' and can felicitously combine with any type of ME (e.g., *es gab Regen, Wind, und Sonne* 'it gave rain, wind, and sun') whereas *fallen* 'fall' can obviously only co-occur with MEs that actually fall. It may prove interesting to investigate this variable further, as it might contribute as a usage factor. Another variable that did not feature as much as anticipated is the difference between main clauses and subordinate clauses, as well as the specific types of subordinate clauses. I mostly focused on *when*/*if* clauses, since these express conditionality and existential/generic conditions or irrealis situations, which means they often refer to non-actualized MEs. Nevertheless, I coded the data for all types of subordinators, which may be worthwhile exploring in the future.

Chapter 4 left some room to explore agentives further, be it sources like clouds or the sky or other weather gods/patron saints. Causal agents like *Frau Holle* 'Mother Hulda' and *Petrus* 'Saint Peter' are certainly not weather concepts that are transparent especially with regard to learners. The same is true for many of the weather compounds, like *Fritz-Walter-Wetter* 'Fritz Walter weather' or *Kaiserwetter* 'Emperor weather'. These often collocated with the verb *herrschen* 'reign/rule' or 'to dominate/prevail', which does occur with other MEs, but also frequently with clearly perceptible atmospheric or ambient conditions such as *Temperature* 'temperature', *Hitze* 'heat', *Dunkelheit* 'darkness', *Lärm* 'noise', *Stille* 'silence', *Klima* 'climate', or *Stimmung* 'mood'. These frequent collocates may turn out to present another minor split type construction not taken into account by Eriksen et al. (2010, 2012).

With regard to metaphorical uses, I came across very few examples, let alone accounts, of weather expressions being used figuratively in German. I introduced a few of these in Section 4.4, such as *hereinschneien* 'to turn up unexpectedly' (lit. 'snow in[to]'), but generally only lightly touched upon the figurative use of weather expressions. Clearly not all weather expressions

appear to lend themselves as readily to metaphorical use as the salient precipitate *Hagel* 'hail' (a bodily experience like temperature, i.e., heat and cold, is presumably a better candidate). This is also true "the other way round", meaning that there are many more metaphorical items that refer to *Regen* 'rain', or rather the process of raining, than to other precipitate types. For German, these were largely related to bodily excretion and/or onomatopoeia. Again, this finding reflects similar observations made by Andrason and Visser (2019) for isiXhosa not just regarding onomatopoeia, but also for the isiXhosa predicate *-na* 'rain type of precipitation', which they hypothesize has a connection to the Pro-Bantu lexical item *nị̀*, which meant 'to defecate/urinate' (cf. also Guthrie 1970).

Lastly, to the best of my knowledge, the present dissertation represents the first in-depth study of weather expressions in German in general and one of the few studies to date that employed an approach that is not purely qualitative (the exception is Liégeois's 2021 study of weather reports). However, the analysis is a fairly basic empirical analysis, which relies on frequencies and proportional data. Following up, I would like to explore a Hierarchical Configurational Frequency Analysis (HCFA) to see what factors and combinations of factors are statistically significant.

5.5 Implications for the lexicon-syntax interface

The crux of this dissertation has been in exploring the rather sensitive interplay between lexis and syntax in a very constricted expressive domain, that is, of intransitive clauses about the weather. The primary focus has been on the interaction between three construction types and the four major precipitate types in German. The data analysis in Chapter 3, especially, has shown how each of the three major constructions appears with a host of construction-specific, but extra-propositional, contextual factors. Following Langacker's concept of a profile-base relationship, where the focus of the attention is the profile, the PREDICATE p-encoding constructions across the board (regardless of precipitate) tend to profile the process of the ME itself in the very reduced sentence structure that is the intransitive clause. This finding is particularly pertinent where the utterance additionally contains a quasi-auxiliary such as *anfangen* 'start' or *aufhören* 'stop', which elevates the incipient or terminative phase of the process in particular. Accordingly, the PREDICATE p-encoding construction data seem very

immediate and the speaker is often directly affected by the event. This was furthermore corroborated by the predominance of present tense instances across all four precipitates and the frequent association with overt markers of stance-taking. In the ARGUMENT p-encoding constructions, it is the reified process, i.e., the 'Thing' or the FIGURE of the precipitate that is semantically heavy and thus profiled. Accordingly, in both types of ARGUMENT p-encoding constructions, the FIGURE is habitually described by referring to the quantity in which the 'thing' comes down. Coincidentally, for the [*geben*+ARG_{PREC}] constructions, this quantity is often zero. Both ARGUMENT p-encoding constructions are also frequently employed to refer to future events, the [*geben*+ARG_{PREC}] construction in particular, which may also explain the frequent association of both constructions with genre type weather reports. The [*fallen*+ARG_{PREC}] constructions generally occur in more formal contexts where stance is expressed infrequently. Other than weather reports this includes the literary genre and meteorological observations about how weather works and in describing the average weather of a region, as is done in travel reports. The diversity within these ostensibly unassuming intransitives is quite intriguing. Fluent speakers, who are frequently exposed to these multi-word sequences, will have internalized the different co-occurrence patterns, even if they cannot fully introspect about when and how to use the different constructional/precipitate pairings.

5.6 Summary and conclusion

Ronald Langacker has famously claimed in cognitive linguistic circles that a language is a "structured inventory of conventional units" (Langacker 1991: 15). Consequently, the basic unit of lexico-grammatical analysis is taken to be a construction, a specific pairing of form and meaning. Constructions can vary in size, nature, and specificity. From a constructional approach, all constructions should potentially be afforded the same importance and analytical attention, be they more recurring and "core" or more idiomatic and "peripheral". Formal accounts and traditional typologies generally tend to focus on morphological and syntactic structure and have less regard for the lexical items that occupy specific "slots" in a phrase or clause. Conversely, lexical semantic accounts often focus primarily on the meaning of lexical items under consideration, neglecting their syntactic context. In subscribing to a cognitive linguistics approach, the analysis presented in this dissertation follows the assumption that grammar and lexicon should not be regarded as separate entities, but must be studied in tandem.

In this study, I have been interested in how a ubiquitous event like the weather is routinely expressed or verbalized in the German language by taking into account both the syntactic frames of the weather constructions as well as the lexical elements that fill those frames. This interest was sparked by a typological study on meteorological expressions (Eriksen et al. 2010), which, among other things, made generalizations about how German encodes precipitation events and, as a native speaker of German, made me instantly consider all types of constructions and expressions I have encountered when talking about the weather. This led to a consideration of different types of encoding precipitation events grounded in a corpus-based approach (using the *deTenTen13* on *SE* and available corpora on the *DWDS*), which revealed even more constructions and ways to speak about the weather in general. The data revealed how remarkably varied—and unruly—weather expressions are, not just cross-linguistically, but also within a single language (German) and across the same sub-domain (such as precipitation). Presumably, this variety and messiness can be attributed to the diffuse nature of the precipitate participant or rather the lack of distinct participants, which poses a challenging lexicalization task for speakers, while tapping into mechanisms of construal and patterns of conventionalization. The perceived impossibility of separating the setting from the participant or the figure from the process itself, usually leads to a dichotomy of one semantically heavy element and one semantically light element (where one is the argument and the other the predicate or vice versa) coming together to form an unmarked, grammatical sentence, mostly in the form of an intransitive clause in German. Additionally, there is the option of argument and the predicate expressing the same semantically heavy information or sharing in expressing different facets of the same event.

The data analysis in Chapters 3 showed that each profile, i.e., each unmarked construction in combination with a specific precipitate type, contained some degree of idiosyncratic variation where sometimes there was overlap between constructions independent of the precipitate type and sometimes there was disparity depending on the precipitate type. In brief, not every precipitate type readily entered into the same construction type and not every construction type hosted the same precipitate type. Furthermore, each [construction x precipitate] interaction occurred with a vast array of often very specific contextual factors. Accordingly, speakers appear to alternate between conventionalized ways of speaking about the weather and the wide variation afforded by a host of available modifiers or reconcile both. These extra-propositional modifiers do not just influence the choice of construction, but also

differ with regard to the precipitate type they allow. Basing an analysis on natural data requires the consideration of subjective factors, such as semantic prosody and stance-taking, where the latter (in contrast to the other non-spatial modifier variables that I coded for) are not part of the semantic frame of the precipitation event despite being habitually employed in at least two of the ME constructions (i.e., the PRED-Cx and the ARG-Cx with *geben* 'give').

Furthermore, the analyses in Chapter 4 showcased several highly marked options to talk about weather events, some of which were culturally specific as they are founded in distinct items like popular songs, fairy tales, and events with such prominence that they have entered the vernacular. These items are not transparent and thus highly idiomatic, which makes them impossible to integrate into a typology of languages. Nevertheless, they are a frequent part of conventionalized language use and as such must feature in a comprehensive account of the German weather "constructicon". This dissertation represents a first attempt at an empirical and comprehensive account of German weather constructions and, hopefully, makes a contribution to a growing roster of languages being explored with regard to their weather expression inventory.

References

Abdulrahim, Dana. 2013. *A corpus study of basic motion verbs in Modern Standard Arabic*. Edmonton: University of Alberta dissertation. https://doi.org/10.7939/R35T12 (last accessed March 20, 2020).

Andrason, Alexander. 2019. Weather in Polish: A contribution to the typology of meteorological constructions. *Studia Linguistica* 73(1), 66–105. DOI: 10.1111/stul.12091

Andrason, Alexander & Marianne Visser. 2017. Cognate objects of weather verbs in African languages of South Africa – From synchronic variation to a grammaticalization path. *Stellenbosch Papers in Linguistics* 48, 151–160. DOI: 10.5774/48-0-287

Andrason, Alexander & Marianna Visser. 2019. Precipitation constructions in isiXhosa. *South African Journal of African Languages*, 39(1), 16–28. DOI: 10.1080/02572117.2019.1572307

Arppe, Antti. 2008. *Univariate, bivariate and multivariate methods in corpus-based lexicography – A study of synonymy*. PhD dissertation, University of Helsinki.

Arppe, Antti, Gaëtanelle Gilquin, Dylan Glynn, Martin Hilpert & Arne Zeschel. 2010. Cognitive corpus linguistics: Five points of debate on current theory and methodology. *Corpora* 5(1), 1–27.

Arppe, Antti & Juhani Järvikivi. 2007. Every method counts – Combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3(2), 131–159.

Aske, Jon 1989. Path predicates in English and Spanish. A closer look. *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society*, 1–14.

Atkins, B. T. S. 1987. Semantic ID tags: Corpus evidence for dictionary senses. *Proceedings of the Third Annual Conference of the UW Centre for the New Oxford English Dictionary*, 17–36.

Austin, Peter K. 1982. Transitivity and cognate objects in Australian languages. In Paul J. Hopper & Sandra A. Thompson (eds.), *Studies in Transitivity*, 37–47. Cambridge: Academic Press.

Barrie, Michael & Bettina Spreng. 2009. Noun incorporation and the progressive in German. *Lingua* 119(2), 374–388.

Bauer, Brigitte 2000. *Archaic syntax in Indo-European: The spread of transitivity in Latin and French*. Berlin & New York: Mouton de Gruyter.

Beavers, John. 2008. On the nature of goal marking and delimitation: Evidence from Japanese. *Journal of Linguistics* 44(2), 283–316.

Benincà, Paola, Guglielmo Cinque. 1992. Sur l'ambiguité structurale des verbes météorologiques en italien [‘On the structural ambiguity of weather verbs in Italian’]. In Liliane Tasmoswki & Anne Zribi Hertz (eds.) *Hommages à Nicolas Ruwet: De la musique à la linguistique* [‘Tributes to Nicolas Ruwet: From music to linguistics’], 155–162. Ghent: Communication & Cognition. The Linguistic Electronic Archive at Ca' Foscari University in Venice, Italy. http://lear.unive.it/jspui/handle/11707/124 (last accessed October 13, 2019).

Bergen, Benjamin & Nancy Chang. 2005. Embodied construction grammar in simulation-based language understanding. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction grammars: Cognitive grounding and theoretical extensions*, 147–190. Amsterdam: John Benjamins.

Biber, Douglas. 1993. Representativeness in corpus design. *Literary and Linguistic Computing* 8(4), 243–257. DOI:10.1093/llc/8.4.243

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *The Longman grammar of spoken and written English*. London: Longman

Biber, Douglas. 2006. *University language: A corpus-based study of spoken and written registers* (Studies in Corpus Linguistics 23). Amsterdam & Philadelphia: John Benjamins. OCLC: ocm65978719.

Biber, Douglas, S. Conrad, R. Reppen (eds.). 1998. *Corpus linguistics: Investigating language structure and use.* Cambridge, New York: Cambridge University Press.

Bleotu, Adina Camelia. 2012. Why does it always rain on me? On weather verbs. *Proceedings of the First Central European Conference in Linguistics for Postgraduate Students*, 59–81.

Bleotu, Adina Camelia. 2013. There is a (light) verb that sometimes goes out in weather verbs. In Irina Windhaber & Peter Anreiter (eds.), *Proceedings of the 4th Austrian Students' Conference of Linguistics*, 47–63. Newcastle upon Tyne: Cambridge Scholars Publishing.

Boas, Hans C. 2013. Cognitive construction grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of construction grammar*, 233–254. Oxford: Oxford University Press.

Bolinger, Dwight. 1973. Ambient *it* is meaningful too. *Journal of Linguistics* 9(2), 261–270.

Bolinger, Dwight. 1977. *Meaning and form*. London: Longman.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald Baayen. 2007. Predicting the dative alternation. In G. Boume, I. Kraemer & J. Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.

Bresnan, Joan & Tatiana Nikitina. 2009. The gradience of the dative alternation. In L. Uyechi & L.H. Wee (eds.), *Reality exploration and discovery: Pattern interaction in language and life*, 161–184. Stanford: CSLI Publications.

Brinker, Klaus, Hermann Cölfen, Steffen Pappert. 2018. *Linguistische Textanalyse: Eine Einführung in Grundbegriffe und Methoden* [Linguistic text analysis: An introduction to basic concepts and methods]. Berlin: Erich Schmidt Verlag.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4), 529–551.

Bybee, Joan. 2010. Language, usage, and cognition. Cambridge: Cambridge University Press.

Bybee, Joan & Clay Beckner. 2009. Usage-based theory. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis* (1st edn.), 827–855. Oxford, UK: Oxford University Press.

Canisius, Theodor. 1867. *Abraham Lincoln: Historisches Charakterbild* [Abraham Lincoln: Historical Character Portrait]. Berlin: Abenheim'sche Verlagsbuchhandlung. (last accessed on Google Books on November 12, 2019)

Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. (2nd revised edn.). Chicago: University of Chicago Press.

Croft, William. 2001. *Radical construction grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

de Cuypere, Ludovic, Evelyn de Coster & Kristof Baten. 2014. The acquisition of the English dative alternation by Russian foreign language learners. *Phrasis: Studies in Language and Literature*, 187–212. Ghent: Academic Press.

Darden, B. 1973. What rains. *Linguistic Inquiry* 4(4), 523–526.

Delisle, H.H. 1986. *Am*/*Beim* progressive constructions in German. In M.C. Marino & L.A. Perez (eds.), *Proceedings of the 12th LACUS Forum*, 1985, 215–224.

Diessel, Holger. 2007. Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology* 25, 108–127.

Diessel, Holger & Martin Hilpert. 2016. Frequency effects in grammar. In Mark Aronoff (ed.), *Oxford research encyclopedia of linguistics.* New York: Oxford University Press. DOI: 10.1093/acrefore/9780199384655.013.120. https://oxfordre.com/linguistics (last accessed November 16, 2019)

Divjak, Dagmar S. & Stefan Th. Gries. 2006. Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2, 23–60.

Divjak, Dagmar S. & Stefan Th. Gries. 2009. Corpus-based cognitive semantics: A contrastive study of phasal verbs in English and Russian. In K. Dziwirek & B. Lewandowska-Tomaszczyk (eds.), *Studies in cognitive corpus linguistics*, 273–296. Frankfurt am Main: Peter Lang.

Dixon, R.M.W. 1994. *Ergativity* (Cambridge Studies in Linguistics 69). Cambridge: Cambridge University Press.

Donaldson, Bruce. 2007. *German: An essential grammar*. New York: Routledge.

Duden. 2006. *Die Grammatik* [The grammar]. Mannheim: Dudenverlag, 2006.

Durrell, Martin. 2011. *Hammer's German grammar and grammar usage* (5th edn). London & New York: Routledge.

Ebert, Karen H. 1996. Progressive aspect in German and Dutch. *Interdisciplinary Journal of Germanic Linguistics and Semiotic Analysis* 1, 41–62.

Ebert, Karen. 2000. Progressive markers in Germanic languages. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe*, 605–653. Berlin & New York: Walter de Gruyter.

Elspaß, Stephan & Robert Möller. 2003. *Atlas zur deutschen Alltagssprache* (AdA) [Atlas of everyday German language]. Online: http://www.atlas-alltagssprache.de

Eriksen, Pål, Seppo Kittilä & Leena Kolehmainen. 2010. The linguistics of weather. Cross-linguistic patterns of meteorological expressions. *Studies in Language* 34(3), 565–601.

Eriksen, Pål, Seppo Kittilä & Leena Kolehmainen. 2012. Weather and language. *Language and Linguistics Compass* 6(6), 383–402.

Eriksen, Pål, Seppo Kittilä & Leena Kolehmainen. 2015. The world is raining: Meteorological predicates and their subjects in a typological perspective. In Marja-Liisa Helasvuo & Tuomas Huumo (eds.), *Subjects in constructions – Canonical and non-canonical* (Constructional Approaches to Language 16), 205–228. Amsterdam & Philadelphia: John Benjamins.

Evans, Nicholas. 2010. Semantic typology. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology,* 504–533. Oxford: Oxford University Press.

Evans, Vyvyan & Melanie Green. 2006. *Cognitive linguistics. An introduction*. Mahwah, NJ: Lawrence Erlbaum.

Evans, Nicholas & David Wilkins. 2000. In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3), 546–592.

Fillmore, Charles J. 1992. "Corpus linguistics" or "Computer-aided armchair linguistics". In Jan Svartvik (ed.), *Directions in Corpus Linguistics, Proceedings of Nobel Symposium 82 Stockholm*, 4–8 August 1991 (Trends in Linguistics. Studies and Monographs [TiLSM] 65), 35–60. Berlin & New York: De Gruyter Mouton. https://is.muni.cz/el/1421/jaro2008/FJ0B738/um/Corpus_linguistics_verze1.pdf (last accessed May 20, 2024)

Firth, John R. 1957. A synopsis of linguistic theory, 1930–1955. In John R. Firth (ed.), *Studies in linguistic analysis. Special volume of the Philological Society*, 1–32. Oxford: Blackwell.

Fried, Mirjam. 2015. Construction grammar. In A. Alexiadou & T. Kiss (eds.), Syntax – Theory and analysis. An international handbook 1 (Handbooks of Linguistics and Communication Science HSK 42.1), 974–1003. Berlin: Mouton de Gruyter.

Fried, Mirjam & Jan-Ola Östman. 2004. Construction grammar: A thumbnail sketch. In Mirjam Fried & Jan-Ola Östman (eds.), *Construction grammar in a cross-language perspective* (Constructional Approaches to Language 2), 11–86. Amsterdam: John Benjamins.

Gárgyán, Gabriella. 2014. *Der am-Progressiv im heutigen Deutsch. Neue Erkenntnisse mit besonderer Hinsicht auf die Sprachgeschichte, die Aspektualität und den kontrastiven Vergleich mit dem Ungarischen*. ['The *am*-progressive in today's German. New insights with particular regard to the history of the language, aspectuality and the contrastive comparison with Hungarian.']. *Networx* 63. Rev. 201312-20. http://www.mediensprache.net/networx/ networx-63.pdf

Garner, Helen. 1992. *Cosmo Cosmolino*. Ringwood, Vic: McPhee Gribble.

Geyken, Alexander.2007. *The DWDS corpus: A reference corpus for the German language of the 20th century.* In Christiane Fellbaum (ed), *Collocations and idioms: Linguistic, lexicographic, and computational* aspects, 23–41. London: Bloomsbury Academic. www.dwds.de/dwds_static/publications/text/geyken_2007_kerncorpus.pdf (last accessed June 4, 2024).

Givón, Talmy. 2001. *Syntax: An introduction*. Amsterdam & Philadelphia: John Benjamins.

Goldberg, Adele. 1995. *Construction grammar. A construction grammar approach to argument structure*. Chicago & London: The University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at work: The nature of generalizations in language*. Oxford: Oxford University Press.

Gries, Stefan Th. 2000. Multifactorial analysis in corpus linguistics: The case of particle placement. Ph.D. dissertation, University of Hamburg.

Gries, Stefan Th. 2003. *Multifactorial analysis in corpus linguistics: A study of particle placement*. London & New York: Continuum Press.

Gries, Stefan Th. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34, 365–399.

Gries, Stefan Th. 2012. Corpus linguistics, theoretical linguistics, and cognitive/ psycholinguistics: Towards more and more fruitful exchanges. In Joybrato Mukherjee & Magnus Huber (eds.), *Corpus linguistics and variation in English. Theory and description* (Language and Computers 75), 41–63. Leiden: Brill.

Gries, Stefan Th. 2012a. Behavioral profiles. A fine grained and quantitative approach in corpus-based lexical semantics. In Gary Libben, Gonia Jarema, Chris Westbury (eds.),

Methodological and analytic frontiers in lexical research, 57–80. Amsterdam & New York: John Benjamins.
www.stgries.info/research/2012_STG_BehavProf_MethAnalFrontLexRes.pdf

Gries, Stefan Th. 2013. Corpus Linguistics: Quantitative methods. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Hoboken, NJ: Blackwell Publishing.
DOI: 10.1002/9781405198431.wbeal0258

Gries, Stefan Th. 2014. Corpus and quantitative methods. In John R. Taylor & Jeanette Littlemore (eds.), *The Bloomsbury companion to cognitive linguistics*, 279–300. London, UK: Bloomsbury Academic. DOI: 10.5040/9781472593689.ch-016
https://pdfs.semanticscholar.org/1089/074625f1b34472786453aed3bda2eab71aff.pdf

Gries, Stefan Th. 2015. The most under-used statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora* 10(1), 95–125. DOI: 10.3366/cor.2015.0068

Gries, Stefan Th. 2015a. Quantitative corpus approaches to linguistic analysis: Seven or eight levels of resolution and the lessons they teach us. In Irma Taavitsainen, Merja Kytö, Claudia Claridge & Jeremy Smith (eds.), *Developments in English. Expanding electronic evidence* (Studies in English Language), 29–47. Cambridge: Cambridge University Press.

Gries, Stefan Th. & Dagmar Divjak (eds.). 2012. *Frequency effects in language learning and processing* (Trends in Linguistics Studies and Monographs 244.1). Berlin & Boston: De Gruyter Mouton. ProQuest Ebook Central, https://ebookcentral.proquest.com/lib/ualberta/detail.action?docID=893502.

Gries, Stefan Th. & John Newman. 2013. Creating and using corpora. In Robert J. Podesva & Devyani Sharma (eds.), *Research methods in linguistics*, 257–287. Cambridge: Cambridge University Press.

Gries, Stefan Th. & N. Otani. 2010. Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal* 34, 121–150.

Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1), 97–129.

Gries, Stefan Th. & Anatol Stefanowitsch (eds). 2006. *Corpora in cognitive linguistics: Corpus-based approaches to syntax and lexis.* Berlin: Mouton de Gruyter.

Grimm, Jacob & Wilhelm Grimm. 1963 [1812]. *Kinder- und Hausmärchen* [Children's and Household Tales], *Frau Holle* [Mother Hulda] #024. Darmstadt: Wissenschaftliche Buchgesellschaft.

Grimm, Jacob & Wilhelm Grimm. 1854–1961. *Deutsches Wörterbuch* [German dictionary], *Erstbearbeitung* [1st edn.] *Digitalisierte Version im Digitalen Wörterbuch der deutschen Sprache* [Digitized version in the Digital Dictionary of the German Language]. dwds.de

Guilquin, Gaëtanelle. 2010. Corpus, cognition and causative constructions (Studies in Corpus Linguistics 39). Amsterdam: John Benjamins.

Guse, Laura. 2022. Source-goal asymmetry in German: A corpus study comparing intentional and non-intentional motion events in the German language. In Laure Sarda & Benjamin Fagard (eds.), *Neglected aspects of motion-event description. Deixis, asymmetries, constructions* (Human Cognitive Processing 72), 173–185. Amsterdam: John Benjamins.

Guthrie M. 1970. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages*. Vols. 1–4. Farnborough: Gregg Press.

Hale, Ken & Samuel J. Keyser. 2002. *Prolegomenon to a theory of argument structure* [Linguistic Inquiry Monographs]. Cambridge, MA: The MIT Press.

Hasko, Victoria. 2012. Qualitative corpus analysis. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics.* Hoboken, NJ: Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal0974 (last accessed August 29, 2024).

Hartmann, Stefan. 2023. Open corpus linguistics – Or how to overcome common problems in dealing with corpus data by adopting open research practices. (preprint) To appear in Mark Kaunisto & Marko Schilk (eds.) *Challenges in corpus linguistics. Rethinking corpus compilation and analysis* [Studies in Corpus Linguistics 118], 89–105. Amsterdam: John Benjamins. https://doi.org/10.31234/osf.io/ywg5v

Hayle, Stephen. 2011. *On the null subjects of Chinese weather verbs*. http://www.gmu.edu/org/lingclub/WP/texts/2_Hayle.pdf (last accessed October 13, 2019)

Heinrich, Claudia. 2014. Raining on the typological parade. Because the weather can be chaotic and unpredictable. Generals paper. University of Alberta.

Heneß, Gottlieb. 1884. *Der neue Leitfaden beim Unterricht in der deutschen Sprache* [New guide for instruction in the German language]. New York: Henry Holt and Company.

Hoffmann, Thomas & Graeme Trousdale (eds.). 2013. *The Oxford handbook of construction grammar*. Oxford: Oxford University Press.

Hopper, Paul J. & Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language* 56, 251–299.

Horita, Yuko. 1996. English cognate object constructions and their transitivity. *English Linguistics* 13, 221–247. DOI:10.9793/elsj1984.13.221

Huang, Chu-Ren, Sicong Dong, Yike Yang & He Ren. 2021. From language to meteorology: Kinesis in weather events and weather verbs across Sinitic languages. *Palgrave Communications* 8(1), 1–13. London: Palgrave Macmillan.

Hunston, Susan. 2002. *Corpora in applied linguistics* (Cambridge Applied Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139524773

Hunston, Susan & Gill Francis. 2000. *Pattern grammar. A corpus-driven approach to the lexical grammar of English* (Studies in Corpus Linguistics 4). Amsterdam: John Benjamins.

Ikegami, Yoshihiko. 1987. 'Source' vs 'goal': A case of linguistic dissymmetry. In René Dirven & Günter Radden (eds.), *Concepts of case*, 122–146. Tübingen: Gunter Narr.

Jackendoff, Ray S. 1983 *Semantics and cognition* (Current Studies in Linguistics Series). Cambridge, MA: The MIT Press.

Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý and Vít Suchomel. 2013. The TenTen corpus family. In *Proceedings of the 7th International Corpus Linguistics Conference* (CL 2013), 125–127, United Kingdom, July 2013.

Janda, Laura A. (ed.). 2013. *Cognitive linguistics. The quantitative turn. The essential reader*. Berlin: De Gruyter Mouton.

Jessen, Christan, Volker Stahl, Erik Eggers & Johann-Günter Schlüper (eds.) 2003. *Fußballweltmeisterschaft 1954 Schweiz. Das Wunder von Bern* [Football World Cup 1954 Switzerland. The Miracle of Bern]. Kassel: AGON Sportverlag.

Juvonen, Päivi & Maria Koptjevskaja-Tamm. 2016. the lexical typology of semantic shifts [Cognititve Linguistics Research 58]. Berlin & New York: De Gruyter Mouton.

Karagjosova, Elena. 2003. Modal particles and the common ground: Meaning and functions of German *ja*, *doch*, *eben*/*halt* and *auch*. In P. Kühnlein, H. Rieser & H. Zeevat (eds.), *Perspectives on dialogue in the new millenium*, 335–349. Amsterdam: John Benjamins. DOI: 10.1075/pbns.114.19kar

Keenan, Edward L. 1976. Towards a universal definition of "subject". In Charles N. Li (ed.), *Subject and topic*, 303–333. New York: Academic Press.

Kilgarriff, Adam & Gregory Grefenstette. 2003. Introduction to the special issue on the Web as corpus. *Computational Linguistics* 29, 333–347.

Koch, Peter. 2001. Lexical typology from a cognitive and linguistic point of view. In Martin Haspelmath, Ekkehard Koenig, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook* – Vol. 2, 1142–1178. Berlin & New York: Walter De Gruyter.

Koch, Peter. 2008. Cognitive onomasiology and lexical change: Around the eye. In Martine Vanhove (ed.), *From polysemy to semantic change: Towards a typology of lexical semantic associations*, 107–137. Amsterdam & Philadelphia: John Benjamins.

Koch, Peter & Wulf Oesterreiche. 1985. *Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte* [The language of proximity – The language of distance. Orality and scribality in the area of tension between language theory and language history]. *Romanistisches Jahrbuch* [Romanistic Yearbook] 35: 15–43.

Koch, Peter & Wulf Oesterreicher. 1994. *Schriftlichkeit und Sprache* [Scribality and writing]. In Hartmut Günther & Otto Ludwig (eds.), *Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung* [Scribality and writing. An interdisciplinary handbook of international research] Vol. 1, 587–604. Berlin & New York: de Gruyter.

Koch Peter & Wulf Oesterreicher. 2012/1985. Language of immediacy - language of distance: Orality and literacy from the perspective of language theory and linguistic history. In Claudia Lange, Beatrix Weber, & Göran Wolf (eds), *Communicative spaces. Variation, contact, and change: Papers in honour of Ursula Schaefer*, 441–473. Frankfurt am Main: Peter Lang. 441-473. DOI: 10.15496/ publikation-20415. [Transl. of Koch and Oesterreicher 1985 by Franz H. Bäuml and Ursula Schaefer]. DOI: 10.15496/ publikation-20415

Kolehmainen, Leena. 2008. *Es schneit wattebauschgroße Schneeflocken*. *Valenz, Proto-Rollen und Argumentstruktur-Konstruktionen bei deutschen und finnischen Witterungsverben* [It's snowing cotton ball-sized snowflakes. Valence, proto-roles and argument structure constructions in German and Finnish weather verbs]. In Mona Enell-Nilsson & Tiina Männikkö (eds.), *Erikoiskielet, käännösteoria ja monikielisyys* [Special languages, translation theory and multilingualism]. VAKKIsymposiumi XXVIII, 100–112. Vaasa 8–9 February 2008. Vaasa: Vaasan yliopisto.

Kolehmainen, Leena. 2010. *Fennosaxonische Wettervalenzen. Turbulenzen im Valenzverhalten deutscher und finnischer Witterungsverben* [Fennosaxon weather valences. Turbulence in the valence behavior of German and Finnish weather verbs]. In Leena Kolehmainen, Hartmut E. H. Lenk & Annikki Liimatainen (eds.), *Infinite Kontrastive Hypothesen. Beiträge des Festsymposiums zum 60. Geburtstag von Irma Hyvärinen* [Infinite contrastive hypotheses. Contributions to the symposium celebrating Irma Hyvärinen's 60th birthday], 59–84. Frankfurt am Main: Peter Lang.

Kollmann, Karl. 2012. *Frau Holle und das Meißnerland: einem Mythos auf der Spur* [Mother Hulda and the Meissner lands: Tracking down a myth]. In Stefan Forbert (ed.). *Historische Gesellschaft des Werralandes und des Werratalvereins*. [Historical Society of the Werraland and the Werra Valley Association], 2nd edn. Heilbad Heiligenstadt, Thuringia: Cordier.

Koptjevskaja-Tamm, Maria. 2008. Approaching lexical typology. In Martine Vanhove (ed.), *From polysemy to semantic change: A typology of lexical semantic associations*, 3–52. Amsterdam: John Benjamins.

Koptjevskaja-Tamm, Maria. 2012. New directions in lexical typology. *Linguistics* 50(3), 373–394.

Koptjevskaja-Tamm, Maria (ed.). 2015. *The linguistics of temperature*. Amsterdam: John Benjamins.

Koptjevskaja-Tamm, Maria. 2016. "The lexical typology of semantic shifts": an introduction. In Päivi Juvonen & Maria Koptjevskaja-Tamm (eds.), *The lexical typology of semantic shifts* (Cognitive Linguistics Research 58), 1–20. Berlin: De Gruyter Mouton.

Koptjevskaja-Tamm, Maria, Dagmar Divjak & Ekaterina V. Rakhilina. 2010. Aquamotion verbs in Slavic and Germanic: A case study in lexical typology. In Viktoria Driagina-Hasko & Renée Perelmutter (eds.), *New approaches to Slavic verbs of motion*, 315–341. Amsterdam & Philadelphia: John Benjamins.

Koptjevskaja-Tamm, Maria, Ekaterina V. Rakhilina & Martine Vanhove. 2016. The semantics of lexical typology. In Nick Rimer (ed.), *The Routledge handbook of semantics*, 434–454. London & New York: Routledge.

Koptjevskaja-Tamm, Maria, Martine Vanhove & Peter Koch. 2007. Typological approaches to lexical semantics. *Linguistic Typology* 11, 159–185.

Krause, Olaf. 2002. *Progressiv im Deutschen: Eine empirische Untersuchung im Kontrast mit Niederländisch und Englisch* [Progressive in German: An empirical study in contrast with Dutch and English]. Tübingen: Niedermayer.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, George. 1991. Metaphor and war: The metaphor system used to justify war in the Gulf. *Peace Research* 23(2), 25–32. Winnipeg: Canadian Mennonite University.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Lakusta, Laura & Barbara Landau. 2012. Language and memory for motion events: Origins of the asymmetry between source and goal paths. *Cognitive Science* 36(3), 517–544.

Lander, Yuri, Maisak Timur & Ekaterina V. Rakhilina. 2012. Domains of aqua-motion: A case study in lexical typology. In Emile van der Zee & Mila Vulchanova (eds.), *Motion encoding in language and space*, 67–83. Oxford: Oxford University Press.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar: Theoretical prerequisites*, Vol. 1. Stanford: Stanford University Press.

Langacker, Ronald W. 1990. *Concept, image, and symbol*. Berlin & New York: Mouton de Gruyter.

Langacker, Ronald W. 1991. *Foundations of cognitive grammar*, Vol. 2. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 1999. *Grammar and conceptualization* (Cognitive Linguistics Research 14). Berlin & New York: De Gruyter Mouton. doi.org/10.1515/9783110800524

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage based models of language*, 1–63. Stanford, CA: CSLI Publications.

Langacker, Ronald W. 2003. Constructions in cognitive grammar. *English Linguistics* 20(1), 41–83.

Langacker, Ronald W. 2005. Integration, grammaticization, and constructional meaning. In Mirjam Fried & Hans C. Boas (eds.), *Grammatical Constructions: Back to the Roots*, 157–189. Amsterdam: John Benjamins. DOI: doi.org/10.1075/cal.4.11lan

Langacker, Ronald W. 2008. *Cognitive grammar. A basic introduction*. New York: Oxford University Press.

Leech, Geoffrey. 1991. Corpora and theories of linguistic performance. In Jan Svartvik (ed.), *Directions in Corpus Linguistics, Proceedings of Nobel Symposium 82 Stockholm*, 4–8 August 1991 (Trends in Linguistics. Studies and Monographs [TiLSM] 65), 105–128. Berlin & New York: De Gruyter Mouton. DOI: 10.1515/9783110867275.105

Leech, Geoffrey. 2002. The importance of reference corpora. UZI. invited speech given by Prof. Leech at the conference of Hizkuntza-corpusak. Oraina eta geroa (2002-10-24/25). https://uzei.eus/wp-content/uploads/2017/06/06-Geoffrey-LEECH.pdf

Leech, Geoffrey. 2007. New resources or just better old ones. In M. Hundt, N. Nesselhauf & C. Biewer (eds.), *Corpus linguistics and the Web*, 134–149. Amsterdam: Rodopi.

Lehmann, Christian. 1990. Towards lexical typology. In William Croft, Keith Denning & Suzanne Kemmer (eds.), *Studies in typology and diachrony. Papers presented to Joseph H. Greenberg on his 75th birthday*, 161–186. Amsterdam & Philadelphia: John Benjamins.

Lehrer, Adrienne. 1974. *Semantic fields and lexical structure*. Amsterdam: North Holland.

Lehrer, Adrienne. 1992. A theory of vocabulary structure: Retrospectives and prospectives. In Martin Pütz (ed.), *Thirty years of linguistic evolution. Studies in honour of René Dirvén on the occasion of his sixtieth birthday*, 243–256. Amsterdam: John Benjamins.

Levin, Beth. 1993. *Verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Levin, Beth. & Bonnie Krejci 2019. Talking about the weather: Two construals of precipitation events in English. *Glossa: A Journal of General Linguistics* 4(1), 58. https://doi.org/10.5334/gjgl.794

Levinson, Stephen C. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge, UK: Cambridge University Press.

Liégeois, Vince. 2021. *Die fachinterne und fachexterne Kommunikation des Wetters. Ein Quaestio-orientierter lexikalischer Vergleich* The communication of the weather to specialized and lay audiences. A Quaestio-oriented lexical comparison]. *Germanistische Mitteilungen* [Germanistic Communications] 47, 199–238. https://doi.org/10.33675/GM/2021/47/12

Lohmann, Theodor. 1960. *Petrus und der Wettergott* [Saint Peter and the weather god]. *Zeitschrift für Religions- und Geistesgeschichte* [Journal for Religious and Intellectual History] 12(2), 112–136. Leiden, The Netherlands: Brill.

Louw, B. 1993. Irony in the text or insincerity in the writer?—The diagnostic potential of semantic prosodies. In M. Baker, G. Francis & E. Togni-ni-Bonelli (eds.), *Text and Technology: In Honour of John Sinclair*, 157–176. Amsterdam: John Benjamins. https://doi.org/10.1075/z.64.11lou

Luodonpää-Manni, Milla, Esa Penttilä & Johanna Viimaranta (eds.). 2017. *Empirical approaches to cognitive linguistics: Analyzing real-life data*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Maisak, Timur & Ekaterina Rakhilina. 2007. *Glagoly dviženija i naxoždenija v vode: leksičeskie sistemy i semantičeskie parametry* [Verbs of motion and location in water: Lexical systems and semantic parameters]. In Timur Maisak & Ekaterina Rakhilina (eds.), *Glagoly dviženija v vode: leksičeskaja tipologija* [Verbs of motion and location in water: Lexical typology], 664–693. Moscow: Indrik.

Majid, Asifa & Melissa Bowerman (eds.). 2007. Cutting and breaking events: A crosslinguistic perspective. [Special Issue]. *Cognitive Linguistics* 18(2), 133–152.

Majid, Asifa, Nicholas J. Enfield & Miriam van Staden (eds.). 2006. Parts of the body: Cross-linguistic categorization. [Special Issue]. *Language Sciences*, 28(2/3), 137–147.

Mang, Theo and Sunhilt Mang (eds.). 2007. *Der LiederQuell: Die schönsten Volkslieder aus Vergangenheit und Gegenwart; Ursprünge und Singweisen* [The songwell: The most beautiful folk songs from the past and present; origins and singing styles), 1005–1006. Eggolsheim: Dörfler Verlag.

Mantha, Suzanne & Igor Mel'čuk. 1984. *Phénomènes atmosphériques dans le dictionnaire explicatif et combinatoire du français contemporain* (DEC) [Atmospheric phenomena in the explanatory and combinatory dictionary of contemporary French], *Revue québécoise de linguistique*m [Quebec Journal of Linguistics] 13(2), 271–323. [Une version améliorée est publiée dans [an improved version is published in]: Mel'čuk et al. 1988: 41–47, "Le champ lexical "Phénomènes atmosphériques": Formulation des définitions" [The lexical field "Atmospheric phenomena": Formulation of definitions].

McEnery, Tony & Andrew Hardie. 2012. *Corpus linguistics: Method, theory and practice*. Cambridge, UK: Cambridge University Press.

McEnery, Tony & Andrew Wilson. 2001. *Corpus linguistics: An introduction* (2nd edn.). Edinburgh: Edinburgh University Press.

McEnery, Tony, Richard Xiao & Yukio Tono. 2006. Corpus-based language studies: An advanced resource book. London, UK: Routledge.

Meulleman, Machteld & Katia Paykin. 2016. Weather verbs sifted through a motion sieve. In *Contrastive Linguistics* 5, 58–67.

Meulleman, Machteld & N. Stockman. 2013. *La inacusatividad en los verbos meteorológicos en español: Un análisis comparativo de llover y amanecer* [Unaccusativity in meteorological verbs in Spanish: A comparative analysis of rain and dawn]. *Bulletin of Hispanic Studies* 90, 117–132.

Michaelis, Laura. 2013. Sign-based construction grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of construction grammar*, 133–152. Oxford: Oxford University Press.

Miller, David, Vincent DeLuca, Denny Berndt, Michael Iverson & Jason Rothman. 2017. Subtle aspectual differences in the L2 acquisition of German: The case of the present tense and pseudo-progressive *am* + infinitive and *beim* + infinitive constructions. In Elma Blom,

Leonie Cornips & Jeannette Schaeffer (eds.), *Cross-linguistic influence in bilingualism: In honor of Aafke Hulk* (Studies in Bilingualism 52), 207–230. Amsterdam & Philadelphia: John Benjamins.

Moravcsik, Edith A. 2006. *An introduction to syntactic theory. Fundamentals of syntactic analysis*. London & New York: Continuum.

Newman, John. 1996. *Give: A cognitive linguistic study*. Berlin: Mouton de Gruyter.

Newman, John (ed.). 1998. *The linguistics of giving* (Studies in Typological Linguistics 36). Amsterdam & Philadelphia: John Benjamins.

Newman, John. 1998a. The origin of the German *es gibt* construction. In John Newman (ed.), *The linguistics of giving* (Studies in Typological Linguistics 36), 307–325. Amsterdam & Philadelphia: John Benjamins.

Newman, John (ed.). 2002. *The linguistics of sitting, standing and lying* (Typological Studies in Language 51). Amsterdam & Philadelphia: John Benjamins.

Newman, John (ed.). 2009. *The linguistics of eating and drinking* (Typological Studies in Language 84). Amsterdam & Philadelphia: John Benjamins.

Newman, John & Sally Rice. 2001. English SIT, STAND, and LIE in small and large corpora. *ICAME Journal* 25, 109–133.

Newman, John & Sally Rice. 2006. Transitivity schemas of English EAT and DRINK in the BNC. In Stefan Th. Gries & Anatol Stefanowitsch (eds.), *Corpora in cognitive linguistics: Corpus-based approaches to syntax and lexis*, 225–260. Amsterdam & Philadelphia: John Benjamins.

O'Keeffe, Anne & Michael McCarthy. 2010. Historical perspective: What are corpora and how have they evolved? In Anne O'Keeffe & McCarthy, M.(eds.), *The Routledge handbook of corpus linguistics*, 3–13. London: Routledge.

Paul, Hermann. 1992. *Deutsches Wörterbuch. Bedeutungsgeschichte und Aufbau unseres Wortschatzes*, 9. *Aufl*. [German dictionary. History of meaning and structure of our vocabulary, 9th edn.], 516. Tübingen: Max Niemayer.

Paykin, Katia. 2010. *Il pleut des idées reçues* [It's raining preconceived ideas: NP expansions of weather verbs]. *Lingvisticae Investigationes* 33, 253–266.

Pollard, Carl and Ivan A. Sag. 1987. *Information-based syntax and semantics*, Vol. 1: *Fundamentals*. Stanford, CA: CSLI Publications.

Pollard, Carl and Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: Chicago University Press.

Pomikálek, Jan. 2011. Removing boilerplate and duplicate content from web corpora. PhD dissertation, Faculty of Informatics, Masaryk University. Brno: Czech Republic.

van Pottelberge, Jeroen. 2004. *Der* am-*Progressiv: Struktur und parallele Entwicklung in den kontinentalwestgermanischen Sprachen* [The *am*-progressive: Structure and parallel development in the continental West Germanic languages]. Tübingen: Gunter Narr Verlag.

Poznański, Samuel Abraham. 1901. *Das Geniza-Fragment* [The Geniza fragment]. In I. Abrahams & C. G. Montefiore (eds.), *Jewish Quaterly Review* XIII, 746–749. London: Macmillan.

Rakhilina, Ekaterina V. & Tatiana Reznikova. 2014. Doing lexical typology with frames and semantic maps. Basic research program. Working papers. Series: linguistics. WP BRP 18/LNG/2014. Moscow: National Research University HSE.

https://www.hse.ru/data/2015/02/12/1092528126/18LNG2014.pdf (last accessed 24 December 2019).

Reimann. Ariane. 1998. *Die Verlaufsform im Deutschen: entwickelt das Deutsche eine Aspektkorrelation?* [The progressive in German: Does German develop aspectual correlation?]. PhD dissertation, Otto-Friedrich-Universität, Bamberg [Microfiche].

Rice, Sally. 1987. Towards a cognitive model of transitivity. PhD dissertation., UCSD, San Diego.

Rice, Sally. 1988. Unlikely lexical entries. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society* 1988, 202–212.

Rice, Sally. 2002. Figure/Motion/Manner/Path conflation patterns in the Dene Sųłiné verb. Paper presented at the 2002 Athapaskan Language Conference in Fairbanks, Alaska, 16–18 June.

Rice, Sally. 2018. Reflections on documentary corpora. In Bradley McDonnell, Andrea L. Berez-Kroeker & Gary Holton (eds.), *Language documentation & Conservation special publication 15. reflections on language documentation 20 years after Himmelmann 1998*, 157–172. http://hdl.handle.net/10125/24817 (last accessed May 20, 2024)

Rice, Sally & Dorothy Thunder. 2017. Community-based corpus-building: Three case studies. Paper presented at the 3rd International Conference on Language Documentation and Conservation. Honolulu, March 2–5, 2017.

Ruwet, Nicolas. 1986. On weather verbs. Chicago Linguistics Society. *Papers from the General Session at the Twenty-Second Regional Meeting* (CLS22): Part 1. April 17–19, 1986. Chicago, Illinois: Chicago Linguistics Society: 195–215

Ruwet, Nicolas. 1991. *Syntax and Human Experience* (Studies in Contemporary Linguistics). Edited and translated by John Goldsmith. Chicago & London: The University of Chicago Press.

Rychlý, Pavel, 2008. A lexicographer-friendly association score. In Petr Sojka & Aleš Horák (eds.), *Proceedings of Recent Advances in Slavonic Natural Language Processing* (RASLAN 2008), Karlova Studánka, Czech Republic, December 5–7, p. 6–9. Brno: Masaryk University. http://www.fi.muni.cz/usr/sojka/download/raslan2008/13.pdf (last accessed May 20, 2024)

de Saussure, Ferdinand. 2011 [1916] *Cours de linguistique Générale* [Course in general linguistics, Perry Meisel & Haun Saussy (eds.)]. New York: Columbia University Press.

Schmid, Hans-Jörg. 2015. A blueprint of the entrenchment-and-conventionalization model. *Yearbook of the German Cognitive Linguistics Association* 3(1), 3–25.

Schmid, Helmut & Florian Laws. 2008. *Estimation of conditional probabilities with decision trees and an application to fine-grained POS tagging*. https://www.cis.uni-muenchen.de/~schmid/papers/Schmid-Laws.pdf (last accessed July 15, 2019)

Schönefeld, Doris. 1999. Corpus linguistics and cognitivism. *International Journal of Corpus Linguistics* 4(1), 137–171.

Schultze-Berndt, Eva. 2007. On manners and paths of refining Talmy's typology of motion expressions via language documentation. In Peter K. Austin, Oliver Bond, David Nathan (eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory*, 223–233, London: School of Oriental and African Studies (SOAS).

Sinclair, John McHardy. 1991. *Corpus, concordances, collocations*. Oxford: Oxford University Press.

Sinclair, John McHardy. 2004. *Trust the text. Language, discourse, and context*. London & New York: Routledge.

Slobin, Dan I. 1996. Two ways to travel: Verbs of motion in English and Spanish. In Masayoshi Shibatani & Sandra Thompson (eds.), *Grammatical constructions*, 195–219. Oxford: Clarendon Press.

Slobin, Dan I. 1997. *The crosslinguistic study of language acquisition* (Expanding the Contexts 5). New York: Psychology Press.

Slobin, Dan I. 2003. Language and thought online: Cognitive consequences of linguistic relativity. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in mind: Advances in the study of language and thought*, 157–192. Cambridge, MA: The MIT Press.

Slobin, Dan I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In Sven Strömqvist & Ludo Verhoeven (eds*.), Relating events in narrative: Topological and contextual perspectives*, 219–257. Mahwah, NJ: LEA Publishers

Slobin, Dan & Nini Hoiting. 1994. Reference to movement in spoken and signed language: Typological considerations. *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society*, 487–505.

Smith, Michael B. 2002. The polysemy of German *es*, iconicity, and the notion of conceptual distance. *Cognitive Linguistics* 13(1), 67–112.

Spreng Bettina. 2017. The absentive: Time, person, and place deixis. *Actes du congrès annuel de l'Association canadienne de linguistique* 2017 [Proceedings of the 2017 annual conference of the Canadian Linguistic Association].

Stefanowitsch, Anatol. 2006. *Corpus linguistics. A guide to the methodology* (Textbooks in Language Sciences 7). Berlin: Language Sciences Press.

Stefanowitsch, Anatol. 2011. Cognitive linguistics meets the corpus. In Mario Brdar, Stefan Th. Gries & Milena Žic Fuchs (eds.), *Cognitive linguistics: Convergence and expansion* (Human Cognitive Processing 32), 257–290

Stefanowitsch, Anatol. 2020. *Corpus linguistics: A guide to the methodology* (Textbooks in Language Sciences 7). Berlin: Language Science Press. DOI: doi 10.5281/zenodo.3735822

Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2), 209–243.

Stefanowitsch, Anatol & Stefan Th. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1(1), 1–43.

Stefanowitsch, Anatol & Stefan Th. Gries (eds). 2006. *Corpus-based approaches to metaphor and metonymy.* Berlin: Mouton de Gruyter.

Stefanowitsch, Anatol & Ada Rohde. 2004. The goal bias in the encoding of motion events. In Günter Radden & Klaus-Uwe Panther (eds), *Studies in Linguistic Motivation*, 249–268. Berlin: Mouton de Gruyter.

Suchomel, Vít and Jan Pomikálek. 2012. Efficient Web crawling for large text corpora. In Adam Kilgarriff, Serge Sharoff. *Proceedings of the Seventh Web as Corpus Workshop* (WAC7), 39–43. Lyon.

Sullivan, Karen Sorensen. 2007. Grammar in metaphor: A construction grammar account of metaphoric language. PhD dissertation, University of California.

Talmy, Leonard. 1972. Semantic structures in English and Atsugewi. PhD dissertation, University of California.

Talmy, Leonard. 1975. Semantics and syntax of motion. In John P. Kimball (ed.), *Syntax and semantics*, Vol. IV, 181–238. New York: Academic Press.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language typology and syntactic description*, 57–149. New York: Cambridge University Press.

Talmy, Leonard. 1991. Path to realization: A typology of event conflation. *Proceedings of the Berkeley Linguistics Society* 17, 480–519.

Talmy, Leonard. 2000. *Toward a cognitive semantics*, Vol. 1 & 2. Cambridge, MA: MIT Press.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Václav Cvrček, Zuzana Komrsková, David Lukeš, Petra Poukarová, Anna Řehořková, Adrian Jan Zasina & Vladimír Benko. 2020. Comparing Web-crawled and traditional corpora. *Language Resources and Evaluation* 54(3), 713–745. https://doi-org.login.ezproxy.library.ualberta.ca/10.1007/s10579-020-09487-4

Velupillai, Viveka. 2012. *An introduction to linguistic typology*. Amsterdam: John Benjamins.

Verspoor, Marjolijn & René Dirven. 2004. Putting concepts together: Syntax. In René Dirven & Marjolijn Verspoor (eds), *Cognitive exploration of language and linguistics* (Cognitive Linguistics in Practice 1, 2nd edn), 75–100. Amsterdam: John Benjamins.

Viberg, Åke. 2001. Verbs of perception. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher& Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, 1294–1309. Berlin & New York: Walter de Gruyter.

Wander, Karl Friedrich Wilhelm (ed.). 1987 [1870]. *Deutsches Sprichwörter-Lexikon*, 1867–1880, Bd. 2. (Lexicon of German Proverbs, 1867–1880. Vol. 2]. Leipzig: Akademische Verlaggesellschaft Athenaion. www.zeno.org/Wander-1867/A/Kaiserwetter (last accessed April 20, 2024)

Wälchli, Bernard & Michael Cysouw. 2012. Lexical typology through similarity semantics: Toward a semantic map of motion verbs. *Linguistics* 50(3), 671–710.

Wasow, Thomas & Jennifer Arnold. 2005. Intuitions in linguistic argumentation. *Lingua* 114(11), 1481–1496.

Wirsel, C.W. 1854. *Wörterbuch zur Fortbildung in der Sprache für diejenigen Taubstummen welche bereits ihre Ausbildung erlangt haben* [Dictionary for further training in the language for those deaf and mute who have already obtained their training]. Lippstadt: H. Lange's Buchhandlung.

Wynne, Martin (ed.). 2005. *Developing linguistic corpora: A guide to good practice*, produced by Oxbow Books and AHDS. https://dx.doi.org/20.500.14106/dlc

Xiao, Richard. 2010. Corpus creation. In N. Indurkhya, F. Damerau (eds.), *The handbook of natural language processing* (2nd edn.), 147–165. London: CRC Press-Taylor & Francis Group.

Zamazal, Kryštof. 2024. Evaluation of Web page cleaning tool Justext. Bachelor's thesis, Faculty of informatics, Masaryk University. Brno, Czech Republic

Zelinsky-Wibbelt, Cornelia. 1995. Insights from categorization for a mental model of lexicalization. In René Dirven & Johan Vanparys (eds.), *Current approaches to the lexicon: A selection of papers presented at the 18th LAUD symposium*. (Duisburger Arbeiten zur Sprach- und Kulturwissenschaft [Duisburg works on linguistics and cultural studies], Vol. 24), 27–61. Frankfurt am Main: Peter Lang Verlag.

Zlatev and Yangklang 2004. A third way to travel: The place of Thai in motion event typology. In S. Strömqvist & L. Verhoeven (eds.), *Relating events in narrative*, Vol. 2: Typological and contextual perspectives), 159–190. Mahwah, NJ: Lawrence Erlbaum.