UNIVERSITY OF ALBERTA

Effects of plural semantics on linking element alternation in German nominal compounds

COMBINED HONOURS THESIS

BY

Elizabeth Pankratz

ABSTRACT

In approximately one third of German nominal compounds, interfixes called linking elements appear between the compound's constituents. In some cases, these linking elements are homophonous with German's plural markers. Linking elements with plural form may sometimes, but do not always, express plural semantics, and no clear explanation of exactly when plural form and meaning overlap has yet been put forward. The research discussed in this paper helps to fill this gap by performing a corpus study that investigates whether the selection of linking elements that are homophonous with plural markers is conditioned by morphological and/or semantic characteristics of the compound's head noun. This also sheds light on the under-researched role that the head constituent plays in linking element selection. The investigation will show that a semantically plural (i.e. collective) head noun is a significant predictor of the appearance of a linking element with plural form and conclude that the predictable co-occurrence of plural form and meaning in such cases suggests that these linking elements may have plural-morpheme-like properties.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

1 Introduction

The linking elements that appear in German nominal compounds make up a notoriously complicated and irregular system. This paper sheds light on an area of relative regularity in linking element distribution by investigating the extent to which, and under which conditions, linking elements may be associated with plurality within German compounds.

Certain linking elements share their form with German's plural suffixes but do not always convey plural meaning, while in other cases, there is doubtlessly plural meaning but non-plural form. Because of this inconsistent relationship between plural form and plural meaning within compounds, some scholars prefer the sweeping position that no linking elements carry any plural meaning, since no generalizations can be made. Others believe that some linking elements may convey plural meaning, and still others go as far as to suggest that, when plural form and plural meaning do coexist within a compound, it's a matter of internal plural inflection.

However, there is no thorough description in the literature of exactly *when* plural meaning or internal plural inflection may appear within compounds. The present study will contribute to filling in this gap in our knowledge of linking element distribution and will suggest that the most appropriate analysis of the linking elements that have plural form and occur in situations of plural meaning is as plural-morpheme-like linking elements.

Previous studies of linking element distribution have identified a wide variety of factors that determine which linking element appears in which compound (see Section 1.1.2). The present study focuses on the question of whether the selection of linking elements that are homophonic with plural inflectional suffixes is morphologically and/or semantically conditioned. In other words, is a linking element with plural form (e.g. *-e-*, homophonic with German's plural suffix *-e*) more likely than a non-plural linking element (e.g. *-ens-*, not homophonic with any of German's plural suffixes) to show up when the compound is morphologically plural? Or perhaps

when the head of the compound (its second constituent) is semantically plural (i.e. a collective noun)? To answer these questions, a study was performed using a 5-billion-token subset of the DECOW16A web corpus (Schäfer & Bildhauer 2012).

This paper also comments on the less well-known role of the second constituent in the selection of a linking element. The influence of the first constituent on linking element selection is inarguably much stronger (see Section 1.1.2) and consequently more heavily researched, but this study demonstrates that certain characteristics of the head constituent also play a very important role when grammatical number is involved, filling in an under-researched area in the investigation of linking element distribution.

Section 1 will build the theoretical foundation for this study. Section 1.1 briefly explains the complexity of the German linking element system and Section 1.2 discusses the nature of the relationship between linking elements and plurality, including a review and evaluation of the three primary positions adopted in the literature. Section 1.3 explores the morphological and semantic avenues through which plurality may appear within the compound. Identifying these allows for the development of a targeted methodology, which will be detailed in Section 2. I will describe the results in Section 3 and review them in detail in Section 4, concluding with a discussion of what this study means for our understanding of the relationship between plurality and linking elements in German compounding.

1.1 Compounding and linking elements

1.1.1 Anatomy of the German compound word

Languages across the world use the word-building process of COMPOUNDING, which is the fusion

of two (or more) existing lexical elements into a single new lexical unit, called a COMPOUND

WORD (Krott et al. 2007:27). Compounding is a very frequent and highly productive process in

German, where speakers can create not only nominal compounds such as *Handschuh*

(Hand$_N$+Schuh$_N$ 'mitten', lit. 'hand shoe'), but also verbal compounds like *kennenlernen*

(kennen$_V$+lernen$_V$ 'to get to know', lit. 'know learn') and adjectival compounds like *bittersüß*

(bitter$_{ADJ}$+süß$_{ADJ}$ 'bittersweet').

German compounds are right-headed, so the right (i.e. second) constituent is the head of

the compound and determines the part of speech and grammatical characteristics (e.g. gender) of

the entire compound; the first constituent can belong to any part of speech. For instance, *feuerrot*

feuer$_N$+rot$_{ADJ}$ 'fire red' is a noun+adjective adjectival compound. This paper will focus

exclusively on German nominal compounds of the type noun+noun.

At its most basic, German nominal compounding simply involves the concatenation of

two nouns, called CONSTITUENTS, in their basic nominative singular form, which is the default

form when building compounds (Neef 2015:30). (German has four cases: nominative, accusative,

dative, and genitive, and two nominal declension classes: strong and weak. The only case in

which no inflectional morphology is applied in either declension class is the nominative. See

Table 4 in Section 2.2 below for an overview of German nominal inflection morphology.) These

straightforward noun+noun combinations account for approximately 65% to 70% of all nominal compounds in German (Gallmann 1999:177, Libben et al. 2002:23; Krott et al. 2007:29).[1]

Some scholars (e.g. Krott et al. 2007) analyze simple noun+noun compounds as actually containing a zero interfix -Ø-, e.g. *Hand-Ø-Schuh* 'mitten'.[2] There is, to my knowledge, no clear evidence for the existence of a zero interfix; it seems to be largely a matter of taste as well as ease of notation. Primarily for the latter reason (and remaining agnostic as to whether there may actually be a zero interfix present), I will use -Ø- in this paper as a shorthand to indicate that only bare nominative singular forms appear in the compound, as in 1.

(1) Nagel   +   Lack   →   Nagel-Ø-lack
     'nail'   'lacquer'   'nail polish'
     Leder   +   Stuhl   →   Leder-Ø-stuhl
     'leather'   'chair'   'leather chair'

In the other 30% to 35% of compounds, one of several changes to the basic nominative form takes place. Most frequently, this change is interfixation: a segment or syllable is inserted between the two constituents, or more accurately, added to the end of the first constituent (see 7 in section 1.1.2 below). These additional interfixes are typically called LINKING ELEMENTS. The

---

[1] These sources do not state explicitly whether the given proportions refer to types or tokens. Gallmann says compounds "with linking elements make up about 30% … the other 70% are formed without a linking element" (1999:177, my translation), Libben and colleagues state that "German compounding … shows the presence of interfixes between constituents in about 35% of all compounds" (2002:23), and Krott and colleages report that "65% of the noun-noun compounds in the CELEX lexical database" (2007:29) appear without a linking element.

[2] German compounds are represented orthographically as a single word, e.g. <Handschuh>. The hyphenation in this paper was added by me to highlight the compounds' internal structure.

inventory of linking elements that most scholars agree on is *-e-, -er-, -s-, -es-, -n-, -en-, -ns-,* and *-ens-* (Neef 2015:31, Krott et al. 2007:29).

Nübling and Szczepaniak (2013) have proposed that the linking elements *-en-* and *-n-* as well as *-ens-* and *-ns-* are so-called "allo-elements" which can more accurately be described as *-(e)n-* and *-(e)ns-*. The alternation between the two is phonologically conditioned; the syllabic variants are used when the first constituent is monosyllabic or ends on a stressed syllable to create a trochee (German's phonological ideal) and the nonsyllabic variants are used when the first constituent already has a trochaic stress pattern. For example, *-en-* appears in *Frau-en-schuh* 'woman's shoe' while *-n-* appears in *Affe-n-baby* 'baby monkey'. However, this claim has been disputed in the literature, since many counterexamples exist, e.g. *Käse-Ø-ecke* 'cheese wedge', where no linking *-n-* appears (Neef 2015:34). See Nübling & Szepaniak (2013, sections 3.2 and 3.6) for further explanation of their analysis and Neef (2015:33–36) for an argument against it. For this study, I will only focus on surface forms, overlooking any possible allomorphy.

Examples of compounds containing linking elements are shown in 2.

(2) Suppe   +   Topf   →   Suppe-n-topf
     'soup'   'pot'   'soup pot'
     Jahr   +   Zeit   →   Jahr-es-zeit
     'year'   'time'   'season'

Other possible changes to the first constituent's form include the SUBTRACTION of a final schwa as in 3, which will be notated in this paper with an internal N-dash rather than a hyphen; REPLACEMENT of the final segment(s) with either *-s-* or *-en-* as in 4; UMLAUT, i.e. vowel change, as in 5, indicated orthographically with two dots above the affected vowel and underlined

throughout this paper for clarity; and UMLAUT + LINKING ELEMENT as in 6. (These example compounds come from Neef (2015:43), Libben et al. (2002:43), and Dict.cc.)

(3)	subtraction	Kirsche 'cherry'	+	Kuchen 'cake'	→	Kirsch–kuchen 'cherry cake'
		Wolle 'wool'	+	Decke 'blanket'	→	Woll–decke 'wool blanket'
(4)	replacement	Hilfe 'help'	+	Mittel 'means'	→	Hilf-s-mittel 'tool'
		Firma 'company'	+	Chef 'boss'	→	Firm-en-chef 'boss of a company'
(5)	umlaut	Mutter 'mother'	+	Heim 'home'	→	Mütter-heim 'shelter for mothers'
		Vater 'father'	+	Generation 'generation'	→	Väter-generation 'father's generation'
(6)	umlaut + linking element	Gans 'goose'	+	Braten 'roast'	→	Gäns-e-braten 'roast goose'
		Buch 'book'	+	Regal 'shelf'	→	Büch-er-regal 'bookshelf'

A debate within the field of linking element research concerns to what extent these changes in form may correlate to changes in meaning. Generally, the structural changes have no effect on the meaning of the compounds, but this paper will argue that certain linking elements may indeed carry semantic information relating to plurality (see section 1.2).

1.1.2 Linking element distribution in a nutshell

The size of the linking element inventory can be overwhelming when one must choose which linking element (if any) to use in a given compound. The distributional system is also very irregular with only a few areas of predictability, much to the second language learner's chagrin.

This section will very briefly illustrate the complexity of linking element distribution, an area that has been researched extensively in the last twenty years.

Before getting into finer details, it is important to show that the linking element forms a unit together with the first constituent of the compound. This analysis is supported by the fact that formal changes in 3-6 (subtraction, replacement, and umlaut) all affect only the first constituent. The coordination reduction test in 7 also shows clearly that the linking element adheres to the first constituent.

(7)   Kapitänsmützen und Admiralsmützen
       'hats of captains and admirals'
       Kapitän-s- und Admiral-s-mützen
       *Kapitän- und Admiral-s-mützen   (Neef 2015:41)

Furthermore, the linking element remains constant when the second constituent is changed.

(8)   Kapitän   +   Mütze   →   Kapitän-s-mütze
      'captain'   'hat'   'captain's hat'
      Kapitän   +   Stuhl   →   Kapitän-s-stuhl
      'captain'   'chair'   'captain's chair'
      Kapitän   +   Amt   →   Kapitän-s-amt
      'captain'   'duty'   'captaincy'

These examples in 8 also suggest that the choice of linking element may be dependant on the identity of the first constituent, since the different second constituents seem to have no effect on which linking element appears. Various characteristics of the first constituent do indeed determine which linking element is selected, though this is where the system's complexity begins.

Characteristics that inform linking element selection include lexeme (i.e. the linking element is lexicalized as "part" of that first constituent), declension class, grammatical gender, plural semantics, and phonology (Nübling & Szczepaniak 2013:84). Particular linking elements are associated with one or more of these characteristics.

To illustrate, the appearance of *-n-* is frequently correlated with the phonology of the first constituent, because it very often appears after schwa-final nouns like *Löwe* 'lion' as in 9. However, its appearance also is partially related to grammatical gender, since "[i]n the case of [schwa-final] feminine nouns, this rule is almost invariant" (Libben et al. 2002:32), for example after the feminine *Woche* 'week' in 10. A second linking element, *-en-*, is morphologically conditioned by declension class, appearing nearly invariably on first constituents that belong to the weak declension class, e.g. *Held* 'hero'; see 11.

(9)	Löwe 'lion'	+	Göttin 'goddess'	→	Löwe-n-göttin 'lion goddess' *Löwe-göttin *Löwe-s-göttin *Löw–göttin
(10)	Woche 'week'	+	Zeitung 'newspaper'	→	Woche-n-zeitung 'weekly newspaper' *Woche-zeitung *Woche-s-zeitung *Woch–zeitung
(11)	Held 'hero'	+	Tat 'deed'	→	Held-en-tat 'heroic deed' *Held-tat *Held-s-tat

For a very detailed descriptive analysis of linking element distribution, see Fuhrhop (1996), and for an intriguing explanation of why the placement of each linking element is determined by such diverse factors, see Nübling & Szczepaniak (2013).

In 9-11, the acceptable forms have also been contrasted with unacceptable forms containing no linking element, the linking element -*s*-, and subtraction in the case of the schwa-final first constituents. This illustrates that any given first constituent not only prefers a particular linking element but also disallows others.

However, some first constituents appear with more than one linking element. A phenomenon central to this paper's study is LINKING ELEMENT ALTERNATION: the ability of a first constituent to take one linking element in one compound and another in a second, in some cases disallowing the linking element from the other compound. Linking element alternation shown by the first constituent *Tag* 'day' is illustrated in Table 1.

Table 1: Linking element alternation for the first constituent *Tag* 'day'

Linking Element	Sample Compound	Head Constituent in Other Compounds
none (-Ø-)	Tag-Ø-traum 'daydream' *Tag-e-traum *Tag-es-traum	Kindheit-s-traum 'childhood dream' Sieg-es-traum 'dream of victory' (lit. victory dream)
-*e*-	Tag-e-buch 'diary' (lit. days book) *Tag-Ø-buch *Tag-es-buch	Wört-er-buch 'dictionary' (lit. words book) Bund-es-buch 'Book of the Covenant' (lit. covenant book)
-*es*-	Tag-es-licht 'daylight' *Tag-Ø-licht *Tag-e-licht	Abend-Ø-licht 'evening light' Himmel-s-licht 'heavenly light' (lit. heaven light)

The examples in the third column show that this alternation does not occur due to the presence of a particular second constituent. Nothing about *Traum* 'dream', for example, consistently prompts -Ø-, since *Traum* can also be the head of compounds that include other linking elements conditioned by the first constituent. Despite this alternation phenomenon, there

is still much more variability between the second constituent and linking element than between the first constituent and linking element.

Neef (2015) suggested that in such situations of linking element alternation, all linking elements but one are "irregular" and "to be treated as lexicalized" and only one linking element is used productively in new compounds (2015:46). The current study will investigate this claim by examining a factor that might in fact determine productive alternation between *two* linking elements, rather than Neef's theorized *one*: the plurality (semantic or morphological) of the second constituent's head.

In the next section, I will thoroughly explain the relationship between linking elements and grammatical number.

1.2 Linking elements and plurality

German's plural inflection is not as complicated and irregular as linking element distribution, but neither is it straightforward. German has four different plural suffixes, *-e*, *-er*, *-(e)n*, and *-s*, as well as a zero plural and umlaut in the noun's stem (Fehringer 2009:151). *-e* and *-er* also co-occur with umlaut to create certain plural forms.

These plural suffixes should look familiar, as they share their form with several linking elements. The first column in Table 2 lists all the German linking elements and marks those that overlap with the form of German plural suffixes with ×. The second column contains a list of nominative singular nouns, the third column shows these nouns in their plural form, and the fourth column shows them as first constituents in a compound where they appear with the linking element from the first column. This means that in all cases where the linking element is marked by ×, the noun's plural form and the noun in first constituent position plus linking element are homophonic. (The example compounds in Table 2 are from DECOW16A.)

Table 2: Formal correspondence between plural suffixes and plural linking elements

Linking Element		Singular	Plural	Compound
-Ø-		Lehrer 'teacher'	Lehrer-Ø 'teachers'	Lehrer-Ø-diplom 'teaching certificate'
-e-	×	Geschenk 'gift'	Geschenk-e 'gifts'	Geschenk-e-laden 'gift store'
-er-	×	Kind 'child'	Kind-er 'children'	Kind-er-buch 'children's book'
-s-		Anfang 'beginning'	Anfäng-e 'beginnings'	Anfang-s-zeit 'starting time'
-es-		Jahr 'year'	Jahr-e 'years'	Jahr-es-ende 'end of the year'
-n-	×	Katze 'cat'	Katze-n 'cats'	Katze-n-allergie 'cat allergy'
-en-	×	Frau 'woman'	Frau-en 'women'	Frau-en-stimme 'woman's voice'
-ns-		Name 'name'	Name-n 'names'	Name-ns-schutz 'name protection'
-ens-		Herz 'heart'	Herz-en 'hearts'	Herz-ens-drang 'heart's desire'
umlaut	×	Mutter 'mother'	Mütter 'mothers'	Mütter-zentrum 'centre for mothers'
umlaut + *-e-*	×	Hand 'hand'	Händ-e 'hands'	Händ-e-druck 'handshake'
umlaut + *-er-*	×	Buch 'book'	Büch-er 'books'	Büch-er-regal 'bookshelf'

This formal overlap allows us to distinguish two sub-types of linking element: PLURAL and NON-PLURAL LINKING ELEMENTS.

A PLURAL LINKING ELEMENT is a linking element that shares its form with the plural suffix of any given first constituent. For example, the compound *Frau-en-stimme* 'woman's voice' contains a plural linking element, *-en-*, since the plural of *Frau* 'woman' is *Frau-en* 'women', formed with the suffix *-en*. However, *-en-* is not considered a plural linking element when it

appears after a noun whose plural is not formed through the *-en* suffix, e.g. *Medikament-en-allergie*, 'allergy to medication' (lit. medication allergy). The plural of *Medikament* 'medication' is *Medikament-e*, which makes *Medikament*'s plural linking element *-e-* (though it in fact rarely occurs with this linking element). -Ø- is an exception and I will discuss it in detail below.

NON-PLURAL LINKING ELEMENTS are all other linking elements and any other compounding form (e.g. subtraction or replacement) a given first constituent may take besides the linking element that is homophonous with the noun's plural marker. *-en-* in *Medikament-en-allergie* is considered a non-plural linking element.

To summarize: identity as a plural linking element is dependent on the identity of the first constituent; even though *-en-* is a plural linking element for some first constituents like *Frau*, it is not for others like *Medikament*. Since most nouns have only one plural suffix, there is normally a one-to-one relationship between first constituents and plural linking elements, while there can be a one-to-many relationship between first constituents and non-plural linking elements, since any linking element that is not the plural one counts as non-plural. Also, whether plural meaning is associated with a plural linking element is unimportant for these definitions, because they are purely based on homophony of linking elements and plural suffixes.

Why is -Ø- (i.e. no interfixation) considered a non-plural linking element when German nouns can take a -Ø plural suffix? First, since it has no form, it cannot share a form with the similarly immaterial -Ø plural suffix. Second, all nominative singular forms are also uninflected, so for nouns with a zero plural like *Lehrer* 'teacher', there would be no way based on wordform alone to distinguish *Lehrer* 'teacher' from *Lehrer*-Ø 'teachers' within compounds like *Lehrer*-Ø-*diplom* 'teaching certificate'. (The grammatical number of zero-plural nouns is otherwise clearly expressed through determiners – *der Lehrer* 'the$_{\text{MASC:SG}}$ teacher' vs. *die Lehrer*-Ø 'the$_{\text{PL}}$ teachers' – as well as through subject-verb agreement and adjectival inflection. These cues are not

available in a compound, leaving the first constituent underspecified in terms of grammatical number.) For these reasons, -Ø- is included with the non-plural linking elements.

Additionally, observant readers may have noticed that -s-, though doubtlessly a plural suffix of German, was not marked in Table 2. In most dialects of German, -s- is never used as a linking element if the first constituent takes -s as a plural suffix (Fuhrhop 1996:533–534, Wegener 2003:433, Koester et al. 2004:1654), so there are practically no cases of overlap between the -s plural suffix and the -s- linking element. For instance, the plural of *Auto* 'car' is *Auto-s*, but the compound **Auto-s-X* (where X is any second constituent) is unacceptable.

Fehringer (2009) does point out that in colloquial northern German dialects where the -s plural inflection is more widespread, -s- can indeed occur as a linking element after nouns that inflect for plurality with the -s suffix, such as *Mädel* 'girl' (*die*PL *Mädel-s* 'the girls') or *Jung* 'boy' (*die*PL *Jung-s* 'the boys'). In these dialects, compounds like *Mädel-s-treff* 'girls' meet-up' and *Jung-s-party* 'boys' party' can be used. Since this is restricted to certain dialects and registers, though, -s- will not be considered a plural linking element for the purposes of this study.

To further illustrate the concepts of plural and non-linking elements, I will apply them to the three compounds shown in Table 1: *Tag-Ø-traum* 'daydream'*, Tag-e-buch* 'diary'*,* and *Tag-es-licht* 'daylight'. The plural form of *Tag* 'day' is *Tag-e* 'days', which means that -e- (as in *Tag-e-buch*) is the plural linking element for *Tag*. This leaves -Ø- and -es- as two of its non-plural linking elements.

Let's take a step back. What is to be gained by distinguishing plural and non-plural linking elements? Defining linking elements as plural and non-plural creates a helpful tool for researching to what extent plurality may exist within compounds, since the correspondence between plural form (i.e. a plural linking element) and plural meaning within compounds (i.e.

semantic plurality of the first constituent) is frequently commented on and contested in the literature. The next section will illustrate the partial co-occurrence of plural form and meaning within compounds and explore scholarly approaches to this issue.

1.2.1 Existing literature about plurality and linking elements

Despite the formal correspondence, plural linking elements do not universally convey plural meaning. In fact, mismatch between form and meaning occurs frequently and can appear in two different configurations: either there is a plural form but no plural meaning or a plural meaning without plural form. The compounds *Frau-en-stimme* 'woman's voice' and *Bild-Ø-sammlung* 'picture collection' will be used to illustrate this.

We saw above that *Frau-en-stimme* contains a plural linking element *-en-*, but the semantic number of women involved in the compound's semantics is ambiguous (it can also be accurately translated into English as 'female voice'). If plural linking elements always conveyed plural meaning, *Frau-en-stimme* would mean *'the voice(s?) of women', which it does not; thus, there is plural form without plural meaning. The opposite configuration, where plural meaning exists without plural form, is exemplified by *Bild-Ø-sammlung* 'picture collection'. Given the fact that a collection must contain multiple items to be a collection at all, if plural linking elements and plural meaning always went hand in hand, one would expect to see only the form *Bild-er-sammlung* 'picture collection' (lit. pictures collection). Since *Bild-Ø-sammlung* is also a compound that German speakers use, with 169 tokens in my subset of DECOW16A, we can conclude that plural meaning does not necessarily prompt a plural linking element.

Because of this form/meaning mismatch, the literature is divided on how to characterize the overlap in form between plural linking element and plural inflection suffixes. I will briefly explain the three main positions in the literature and show along the way how two of the

compounds we just saw – *Frau-en-stimme* 'woman's voice' and *Bild-er-sammlung* 'picture collection' – would be analyzed differently in each of the three positions.

Position 1: No relation exists between any linking elements and plural semantics.

Neef (2015) takes a staunchly anti-functionalist position. Since linking elements with a plural form cannot be universally associated with plural meaning (i.e. they do not have a function as plural morphemes), he states that linking elements should be thought of as entirely devoid of semantic content (Neef 2015:37–38).

Wegener (2003) argues for the same position based on a diachronic analysis of linking element development. According to her, both plural linking elements and plural suffixes evolved from Old High German along parallel paths from Indo-European stem-building suffixes (Wegener 2003:428). Thus, since one did not lead to the other, the formal overlap is coincidental due to the quirks of linguistic evolution and no plural meaning can be associated with linking elements.

What matters most for supporters of this position is consistency. Since not all linking elements with plural form, (e.g. *-en-* in *Frau-en-stimme*) denote multiple entities of the first constituent's referent (e.g. multiple women), to say that other linking elements can denote multiple entities and express plural meaning would lead to an unreliable analysis. For supporters of this first position, the semantic necessity of having multiple pictures in a collection has nothing to do with the occurrence of *-er-* in *Bild-er-sammlung*. Therefore, both the compounds *Frau-en-stimme* and *Bild-er-sammlung* would be analyzed as containing a linking element which is semantically empty.

Position 2: Some linking elements carry and convey plural meaning.

Researchers who take this position (Gallmann (1999) and Fuhrhop (1996), among others) believe that linking elements are not all semantically empty; some may convey plural meaning when their form matches that of the first constituent's plural inflectional suffix (Gallmann 1999:188, Fuhrhop 1996:534). Applied to our example compounds, this means that *-en-* in *Frau-en-stimme* does not convey any plural meaning, despite the form of the linking element matching that of *Frau*'s plural suffix, while *-er-* in *Bild-er-sammlung* is one of the linking elements that does convey plural meaning (i.e. the first constituent's referent is plural), since both plural form and plural meaning exist in that compound.

Position 3: When a plural linking element co-occurs with plural meaning of the first constituent, it should actually be considered a plural inflectional morpheme that can appear in compounds.

Proponents of this position, e.g. Dressler (1987) and Fehringer (2009), also recognize that plural meaning can sometimes be expressed compound-internally. However, their explanation for the limited correspondence of plural form and plural meaning is more drastic; they make a fundamental terminological and categorical distinction between linking elements and plural inflectional suffixes. According to their analysis, these are two mutually exclusive categories which may *both* appear between compound constituents.

Interfixes with plural form and plural meaning, like *-er-* in *Bild-er-sammlung*, are considered "internal inflectional suffix[es]" (Dressler 1987:75). They are instances of ordinary plural morphology that appear on the compound's non-head constituent (Fehringer 2009:155). Dressler suggests that these are "word-form-based compounds" (1987:75) built from the already inflected form of the first constituent rather than through the combination of two bare nominative singular nouns and a linking element insertion process.

In comparison, interfixes with no plural meaning, regardless of their form, are the only ones considered to be linking elements. *-en-* in *Frau-en-stimme* would be considered an "interfix, i.e. a derivational linking element" (Dressler 1987:75).

The literature around linking elements is overall mostly consistent in terms of vocabulary. However, it is important to be aware of this terminological difference in the meaning of "linking element", because if somebody were to say, "linking elements do not convey plural meaning", the position they represent is in fact unclear. They might be supporting Position 1 (i.e. no plural meaning exists within compounds), but they could also be arguing for Position 3, where linking elements don't convey plural meaning while compound-internal inflectional affixes do.

1.2.2 Evaluation of these positions

Position 1:

Nübling and Szczepaniak (2013:68–74) dispute Wegener's claim on diachronic grounds, stating that "there is no historical evidence" (2013:70) for her analysis that both plural linking elements and plural suffixes evolved from Indo-European stem-building suffixes. They argue that linking elements and plural morphology could not have developed in parallel, since there was a split in the formal development of these in the pre-Old High German period: the primary suffixes were grammaticalized directly into plural markers, but in compounds, they "underwent analogical levelling and formal reduction" (2013:69) to become linking elements. Nübling and Szczepaniak suggest that Wegener did not consider other possible evolutionary factors like analogy in the development of linking elements and that the only basis for her analysis is the non-correspondence of plural form and meaning, rather than historical fact.

More generally, descriptions of linking elements as semantically empty cannot be accepted as completely true, since linking elements can in some cases contribute semantic

information about the first constituent's plurality to the compound, as I will illustrate in the next section.

Position 2:

The third position, though intuitive and straightforward, is unable to explain when linking elements express plural meaning and when they do not. However, because it considers all interfixes to be the same type of thing, namely a linking element, it has the benefit of not introducing more complication (e.g. different types of morphology) into the analysis.

Position 3:

The distinction between internal inflectional suffixes and linking elements is a clean way to explain why some interfixes seem to convey plural meaning and some do not. However, one of its weaknesses is that allowing both inflectional and derivational morphemes to appear between compound constituents is perhaps more complicated than is necessary or even accurate, since it has been shown that plural linking elements are not interpreted – at least in the aural modality – as plural morphemes (Koester et al. 2004). (The perception of linking elements in the aural modality is influenced by prosodic characteristics of compounds. Speakers produce a compound's first constituent more quickly than they would produce a standalone noun, so the listener already knows not to process the linking element as a plural marker (Koester et al. 2004:1662). This may be different in the written modality, however, where these prosodic cues are not present.)

Dressler's suggestion of "word-form-based compounds" (1987:75), i.e. that compounds are built from an already inflected form of the first constituent, is not unprecedented in German. For example, certain adjectival compounds like those in 13 show internal inflection for degree.

(13)   *langfristig*   *lang-fristig*   long-term   Tokens in DECOW16A: 130 406

          kurzfristig   *kurz-fristig*   short-term   Tokens in DECOW16A: 101 490

          längerfristig   *läng-er-fristig*   longer-term   Tokens in DECOW16A: 16 578

          *kürzerfristig   *kürz-er-fristig*   shorter-term   Tokens in DECOW16A: 10

This process is not a productive one, though, since the analogically created form

*kürzerfristig 'shorter-term' is not generally acceptable (though it has been documented 10 times

in DECOW16A).

This position shares its other weakness with Position 2; both suffer from the same

inability to determine *when* linking elements convey plural meaning and when they do not. Since

no satisfactory explanation of this exists in the literature, this paper intends to help fill this gap by

researching under which conditions plural linking elements may be associated with plural

meaning.

1.3 What determines when the first constituent is plural?

This section will identify two sources to which plurality of the first constituent can be

traced back: the first constituent and the head constituent. Sections 1.3.1 and 1.3.2 will focus on

two possible ways the head constituent can "force" plural meaning on the first constituent:

through plural morphology and through its own plural semantics (which we've already

encountered in the example of *Bild-er-sammlung* 'picture collection'). Figure 1 expresses this

relationship visually:

Figure 1:

Sources of plurality of the first constituent

First constituent
(Section 1.3)

Head (second)
constituent

Plural morphology
(Section 1.3.1)

Plural semantics
(Section 1.3.2)
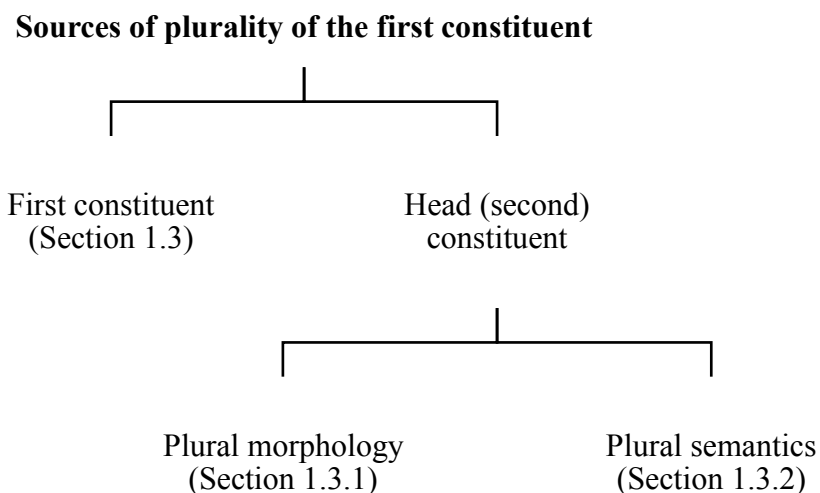
The plural meaning comes from the first constituent when the speaker explicitly chooses

to convey that this constituent is plural. This is exemplified by the minimal pairs in 14 and 15

which show the possible semantic impact of non-plural versus plural linking elements.

Non-Plural Linking Element	**Plural Linking Element**
(14) *Generation-s-konflikt* 'conflict within a generation' (lit. generation conflict)	*Generation-en-konflikt* 'conflict between two generations' (lit. generations conflict)
(15) *Partei-Ø-gruppe* 'faction' (lit. [political] party group)	*Partei-en-gruppe* 'group of parties' (lit. [political] parties group)

It is important to note that, though both plural and non-plural linking elements are

acceptable in the compounds in 14 and 15, not all first constituents in German allow a productive

alternation like this. Alternation based on semantic number seems to be restricted to certain

morphological and phonological conditions. Fuhrhop (1996:534) proposes that plural/non-plural

alternation is productive when the first constituent ends in a feminine-noun-building suffix as in 14, i.e. Germanic suffixes *-heit/-(ig)keit*, *-schaft*, and *-ung*, as well as loan suffixes *-ion* and *-(i)tät*. Nübling and Szczepaniak (2013:77) also note that the alternation may be productive in the case of mono- and polysyllabic feminine nouns with word-final stress and optionally a final consonant, like *Geburt* 'birth' or *Partei* '[political] party', as in 15.

These examples show that a speaker may choose a plural linking element over a non-plural one when he or she wishes to express that the first constituent is plural (and the first constituent allows this type of alternation).

This alternation is based on speaker intention and context and will not be part of this study. Instead, this paper will investigate alternation that is prompted by the head (i.e. second) constituent. Both plural morphology and plural semantics of the head noun could conceivably influence the number semantics of the first constituent, which could lead to the selection of a plural linking element over a non-plural one. These two factors will be explored in more detail in the next two sections.

1.3.1 Plural morphology on the compound's head noun

When a compound's head is inflected in German for number and case, there is generally no effect on the internal structure of the compound, apart from some archaic forms such as *der Hohepriester* 'the high priest' where both constituents inflect, e.g. genitive: *des Hohe-n-priester-s* 'of the high priest' (Dressler 1987:73). Such items, however, are very rare and can also be found written separately (*der Hohe Priester, des Hohen Priesters*); they therefore have much more in common with noun phrases than with compounds.

Also, the fact that the compound has internal structure makes no difference to the inflection of its head. A compound is pluralized the same way its head noun is pluralized when this noun appears alone, as shown below in 16 and 17.

	Singular	**Plural**
(16)	*die*$_{SG:FEM}$ *Stimme*	*die*$_{PL}$ *Stimme-n*
	'the voice'	'the voices'
(17)	*die*$_{SG:FEM}$ *Frau-en-stimme*	*die*$_{PL}$ *Frau-en-stimme-n*
	'the woman's voice'	'the women's voices'

Consider the translation of *die Frau-en-stimme-n* in the second column of 17: 'the women's voices'. Something interesting happens to the meaning of *Frau-en-* when the head noun, *Stimme* 'voice', is pluralized: when there is more than one voice, there must suddenly be more than one woman. In other words, there is a one-to-one relationship between women and voices; a woman has a single voice and cannot have more than one. In this example, the first constituent's semantics has been made explicitly plural through morphological inflection of the head constituent.

This semantic change seems to be something that language users are occasionally sensitive to when constructing compounds. I noticed in compounds from DECOW16A that when the compound's head was inflected for number (i.e. when it was morphologically plural), a plural linking element occasionally appeared or replaced the "usual" non-plural linking element. With *Frau-en-stimme*, the semantic change is not reflected in the form because *Frau* overwhelmingly takes the linking element *-en-*. However, the compound *Mutter-Ø-herz* 'mother's heart' (lit. mother heart) can illustrate this alternation more clearly because its non-plural and plural linking elements are different (*-Ø-* and umlaut, respectively).

The expected plural form with no internal change is *Mutter-Ø-herz-en,* 'mother's hearts' (lit. mother hearts). However, since there is also a one-to-one relationship between mothers and hearts, if heart becomes plural, there must also be more than one mother. So, in addition to the expected plural form, forms such as *Mütter-herz-en* (lit. mothers hearts) appear in DECOW16A.

This pattern, a potential co-occurrence of plural inflection of the compound's head and the selection of a plural linking element over a non-plural one, is one of the two patterns being studied in this paper. The second is a possible co-occurrence of plural linking elements and plural semantics of the compound's head, detailed in the next section.

1.3.2 Plural semantics of the compound's head noun

As alluded to above in the *Bild-Ø-sammlung* 'picture collection' example, there are certain head nouns whose semantics demand the presence of more than one of the first constituent's referent. Consider the examples in 18 and 19. The first column includes compounds with non-plural linking elements and heads with no semantic number effect on the first constituent. The second column lists compounds that have the same first constituent as in the first column, though with a plural linking element and a head noun that forces the first constituent to bear plural meaning. (All examples are from DECOW16A.)

	Head without plural semantics, non-plural linking element	Head with plural semantics, plural linking element
(18)	Lied-Ø-text 'lyrics' (lit. song text)	Lied-er-sammlung 'collection of songs' (lit. songs collection)
	Kind-s-braut 'child bride'	Kind-er-schar 'flock of children' (lit. children flock)

(19) Buch-Ø-binder   Büch-er-wechsel
      'bookbinder'   'book exchange'
                                       (lit. books exchange)

      Stadt-Ø-mauer   Städt-e-wettkampf
      'city wall'   'competition between cities'
                                       (lit. cities competition)

What distinguishes the head nouns in the second column of 18 and 19 from one another is the type of plural semantics they express. In 18, *Sammlung* 'collection' and *Schar* 'flock' are nouns that, even when they are grammatically singular, refer to an entity that has multiple members or segments. A *Lied-er-sammlung* 'song collection' is not a collection if it only contains one song; a *Kind-er-schar* 'flock of children' cannot exist if there is only one child. These nouns have been described as having "lexicalized numerosity" (Nenonen & Niemi 2010:104) and I will call them TRUE COLLECTIVES.

The head nouns in the second column of 19, *Wechsel* 'exchange' and *Wettkampf* 'competition', are abstract nouns that refer to a multi-party action or event. This sort of noun is mostly deverbalized and denotes some sort of interaction between two (or more) objects or parties. When used as a head noun in compounds, the objects or parties that this type of noun requires are multiple instances of the referent of the first constituent. Therefore, *Büch-er-wechsel* is the exchange of multiple books and *Städt-e-wettkampf* is a competition between multiple cities. Since these words describe relations between multiple participants, I will call them PLURAL ACTION COLLECTIVES. For the rest of this paper, however, they will be referred to together with true collectives simply as COLLECTIVE NOUNS or COLLECTIVES.

1.4 Summary

In this section, we have seen that German nouns, when placed in first constituent position in nominal compounds, can undergo several formal changes. The most frequent of these changes is the linking element, which is the interfixation of one of about 10 possible segments or syllables onto the end of the first constituent. The linking element's form is determined based on various characteristics of the first constituent. Some of these linking elements correspond in form to German's plural suffixes, and in many cases (though not always), these plural linking elements co-occur with plural meaning of the first constituent. This partial co-occurrence has been the subject of much discussion in the literature, with some researchers arguing that linking elements cannot carry plural meaning because the correspondence isn't universal and others arguing that they can, though exactly when they do is not clear. Two possible sources of plural semantics which can prompt the alternation between plural and non-plural linking elements were discussed. Plural semantics can come from the first constituent when speakers choose a plural linking element over a non-plural one to highlight the plurality of the first constituent. Plural semantics can also come from morphological or semantic characteristics of the second constituent, which is the focus of this paper's study.

Fuhrhop (1996) states, "In the case of an expressly plural meaning, the corresponding plural inflection is possible as a linking element" (1996:534, my translation). Both morphological plural inflection and semantically collective head nouns can create situations of expressly plural meaning for the first constituent. The aim of this paper is therefore to explore if a plural linking element is chosen over a non-plural one when plural meaning is created through these two avenues.

2 Methodology

As mentioned above, this study was performed using a subset of the DECOW16A web corpus, which contains 5 billion tokens from online texts ranging from forums to official business and university websites, thus including diverse registers.

In short, the study was performed by analyzing samples of compound words gathered from the corpus where the first constituent and linking element were fixed and the head constituent was variable. For each first constituent selected for the analysis, I generated two samples. All compounds in the first (non-plural) sample were composed of that first constituent, any non-plural linking element(s) it occurs with, and any head constituent; the second (plural) sample contained the first constituent, its plural linking element, and any head constituent. Finally, 100 unique and randomly selected tokens in each of these samples were annotated by hand based on morphological and semantic characteristics of the head noun.

In Section 2.1, I will talk about which first constituents I chose for the analysis. Section 2.2 will discuss the process of uniquing the samples and the pros and cons of this procedure and Section 2.3 will explain how the uniqued samples were then annotated.

2.1 First constituent selection

Linking element alternation cannot be studied unless the first constituent is constant between compounds, which is why the first constituent was controlled for and the head constituent was not.

I started by generating lists of the most frequent first constituents for every linking element that can be a plural linking element (which are umlaut, umlaut + *-e-*, umlaut + *-er-*, *-e-*, *-er-*, *-n-*, and *-en-*) The reason for beginning with high-frequency first constituents was to ensure sufficiently large sample sizes. Nenonen and Niemi (2010) note that the generation of samples

based on corpus metrics like frequency should be done with caution, as corpus frequency does not reflect real-life language use (2010:113). However, since the current study is entirely corpus-based, rather than Nenonen and Niemi's experiment with human participants, and because I will narrow down the corpus-generated lists by hand, using corpus-generated token frequency at this early stage of the study does not pose a significant problem for the analysis.

The following nine criteria were used to select suitable first constituents from these frequency lists. To be eligible, first constituents must:

1.	be frequently represented in the corpus (>1000 tokens in total with any linking element and head).

2.	appear in compounds with its plural linking element.

3.	show a reliable alternation (>100 tokens each for plural and non-plural linking elements).

4.	be count nouns.

5.	be concrete nouns.

6.	have only one plural form.

7.	be monolexemic (i.e. not a compound itself).

8.	not be a collective noun.

9.	contribute only its core semantics to the compound or have predictably motivated figurative meaning (which will be methodically excluded in the annotation phase).

For the rest of this section, I will discuss the motivations behind these nine conditions in detail and give examples of the type of words that were disqualified by each.

1. The lower limit of first constituent frequency is defined as 1000 tokens in the corpus, which means that if a first constituent appeared in fewer than 1000 compounds (with any linking element and any head), it was disqualified. For example, *Biest* 'beast' is the 23rd most common first constituent for the linking element *-er-*. However, it appears with its plural linking element *-er-* in 103 compounds (e.g. *Biest-er-bund* 'association of beasts') and with no linking element in only 336 compounds (e.g. *Biest-Ø-rassen* 'beast breeds'). Since only 439 tokens contain this first constituent in the corpus, the sample size is insufficient for this study.

2. A first constituent must regularly occur with its plural linking element in compounds. This is almost always the case, but there are certain exceptions, as we saw in Section 1.2 with *Medikament* 'medication'. Its plural linking element is *-e-*, since it builds its plural with the suffix *-e*, but *Medikament* occurs overwhelmingly frequently with *-en-* instead (in 11462 tokens compared to only 34 with *-e-*); for example, *Medikament-en-allergie* 'allergy to medication'.

To generate the frequency lists, I searched based on all linking elements that *can* be plural linking elements; this doesn't mean they *must* be. When they are not, like in this example, the nouns must be disqualified (unless, of course, they also occur above the 1000-token threshold of their plural linking element, though that was never the case in my data set).

3. To study the alternation phenomenon, it is essential to select only first constituents that show a reliable alternation between the appropriate plural linking element and any non-plural linking element(s), including -Ø-. I quantified "a reliable alternation" to mean that both samples (the plural one and the non-plural one) must be at least 100 tokens in size. In cases of more common linking elements like *-e-*, both samples are generally quite a bit larger, but this cut-off number was set relatively low to still allow the selection of several first constituents from the less common linking elements like umlaut.

For example, *Graben* 'ditch, trench', the 10th most frequent first constituent for umlaut, occurred only 80 times in its plural umlauted form (e.g. *Gräben-system* 'trench system'), though it appeared in 5252 compounds with no linking element (e.g. *Graben-Ø-krieg* 'trench war'). This shows *Graben*'s overwhelming preference for no linking element, i.e. a non-plural form, meaning that it does not have reliable plural/non-plural alternation.

4 and 5. Count nouns and concrete nouns undergo no extreme semantic changes when pluralized, but the same cannot be said for mass and abstract nouns. Even though both mass and abstract nouns generally can take plural morphology, pluralization leads to more dramatic semantic changes than is the case with count and concrete nouns in German as in English. Plural morphology on mass nouns tends to distinguish different types or separate units of something, e.g. *Wein* 'wine' vs. *zwei Weine* 'two glasses of wine, two types of wine', while pluralization of abstract nouns concretizes them: *Erfolg* 'success' becomes *zwei Erfolge* 'two successes', i.e. two instances of success. In these cases, alternation in plural/non-plural form would provoke more extreme semantic changes, rather than simply one vs. multiple instances of the referent, which is the target of this investigation.

Additionally, this criterion leads to the disqualification of a semantic group of nouns: any first constituents denoting animals that are also used to make food and other products, like *Rind* 'beef, cow' and *Lamm* 'lamb', are excluded. Though the creatures themselves are discrete entities, their names primarily have a mass-noun-like usage when appearing as first constituents in compounds, because the compound most often denotes something made from them (e.g. *Rind-s-leder* 'cowhide' or *Lamm-kotelett* 'lamb chop'). Words that do not denote animals but that can also be ingredients with a mass-noun-like usage like *Apfel* 'apple' and *Ei* 'egg' are left in, because they are also used to refer to their discrete entities, i.e. still used consistently in count and concrete noun senses.

6. Next, to keep the analysis straightforward and consistent, eligible first constituents would need to only have one plural form. This is the case for most German nouns, so very few words were disqualified this way, but among them was the highly frequent *Wort* 'word' (the 7th most frequent first constituent for the linking element umlaut + *-er-*). *Wort* has two semantically divergent plurals: *Worte* 'combination of words, saying' and *Wörter* 'many single words'.

7. The exclusion of polylexemic first constituents (i.e. first constituents that are compounds themselves) is so that the present study can focus on compounds with only two constituents, rather than three or more. Consider *Fach-kraft* 'professional' (lit. specialized worker), the fourth most common first constituent generated for the linking element umlaut + *-e-*. Compounds that it appears in, like *Fach-kraft-suche* 'search for professionals' (lit. specialized worker search), have internal hierarchical structure – [[*Fach-kraft*]-*suche*] – which makes them too different from the two-constituent compounds that will be analyzed in the present study.

8. The first constituent must not be a collective noun of either sort, since the potential effect of collective head nouns would be confounded if the first constituent were also a collective. I will illustrate this point using the compound *Familie-n-konflikt* 'family conflict' (lit. families conflict). Because the first constituent *Familie* 'family' (the most frequent first constituent for *-n-*) is a collective noun, it's ambiguous whether the conflict takes place between two families or within one family. This renders the plurality of the first constituent impossible to judge, which is a problem for the analysis.

9. Finally, to avoid figurative meaning and lexicalization effects on the constituent level, first constituents must consistently contribute the same semantics to the compound in both singular and plural forms as they have in their basic nominative form when used alone.

Gast 'guest', the most common first constituent of umlaut + *-e-*, is one of the nouns that this condition eliminates. It contributes its core meaning – the meaning of the noun when used

alone in nominative singular – when used with its plural linking element, like in *Gäst-e-buch* 'guest book'. However, when used with no linking element, it seems to take on a figurative meaning of 'visiting' in compounds like *Gast-Ø-mannschaft* 'visiting team' (lit. guest team) or of 'hosting guests' in *Gast-Ø-familie* 'host family, homestay family' (lit. guest family). This is a problem for the analysis because lexicalized first constituent + linking element units do not allow the same flexibility in their linking element selection. This is the drawback inherent in selecting the most frequent first constituents, because these are more likely to be lexicalized than less frequent, productively used combinations of linking element and first constituent. However, this condition mitigates this effect.

Additionally, there are several first constituents with figurative meaning triggered by the semantics of the compound's head, for example, *Mutter* 'mother'. In most compounds and with both types of linking elements, *Mutter* conveys its core semantics (e.g. *Mutter-Ø-instinkt* 'motherly instinct' and *Mütter-gemeinschaft* 'community of mothers'). However, when the head noun denotes an organization – especially a political one or one run for profit – and no linking element is used, the resulting meaning of the compound is 'an organization from which another one has sprung' (e.g. *Mutter-Ø-konzern* 'mother corporation'). In such situations where the figurative semantics are predictable and can be excluded from the analysis on a case-by-case basis, those tokens will be skipped in the annotation and the first constituent may remain in the analysis.

Based on these nine criteria, suitable first constituents for the present examination were chosen from the frequency lists. The most frequent 12 for each linking element were selected, though for the less common linking elements, fewer were eligible. In total, 69 were chosen as suitable and shown in Table 3 below (though not all of these were analyzed; see Section 2.3).

Table 3: Eligible first constituents sorted by their plural linking element

Umlaut	Umlaut + -e-	Umlaut + -er-	-e-	-er-	-n-	-en-
Mutter 'mother'	Stadt 'city'	Buch 'book'	Hund 'dog'	Kind 'child'	Sonne 'sun'	Frau 'woman'
Vater 'father'	Hand 'hand'	Kraut 'herb'	Spiel 'game'	Mitglied 'member'	Kirche 'church'	Mensch 'human'
Apfel 'apple'	Zahn 'tooth'	Haus 'house'	Gerät 'device'	Bild 'picture'	Kunde 'customer'	Person 'person'
Nagel 'nail'	Ball 'ball'	Bad 'bath(room)'	Film 'movie'	Kleid 'dress'	Straße 'street'	Herr 'mister'
Vogel 'bird'	Stuhl 'chair'	Grab 'grave'	Weg 'path'	Ei 'egg'	Auge 'eye'	Student 'student'
	Baum 'tree'	Rad 'wheel, bicycle'	Meer 'sea'	Lied 'song'	Karte 'card'	Autor 'author'
	Anwalt 'lawyer'	Schloss 'castle'	Geschenk 'gift'	Weib 'woman'	Bauer 'farmer'	Held 'hero'
		Dorf 'town'	Produkt 'product'	Brett 'board'	Blume 'flower'	Ohr 'ear'
		Loch 'hole'	Stück 'piece'	Gespenst 'ghost'	Katze 'cat'	Unterschrift 'signature'
		Wurm 'worm'	Zertifikat 'certificate'	Schwert 'sword'	Maschine 'machine'	Dämon 'demon'
			Brief 'letter'	Gesicht 'face'	Gitarre 'guitar'	Bett 'bed'
			Freund 'friend'		Löwe 'lion'	Burg 'castle'

2.2 Drawing and uniquing samples

I extracted two samples of maximum 2000 tokens each for every first constituent considered for

analysis. (The minimum sample size was 100, corresponding with criterion 2 above.)

As I mentioned already, the first sample was the non-plural sample, including compounds

where the first constituent appeared either without a linking element, with subtraction (common

for schwa-final first constituents), or with any other non-plural linking element(s). For example,

in the first sample of *Kind* 'child' (pl: *Kind-er*), it appeared with no linking element and with the non-plural linking elements *-s-* and *-es-*. The second sample was the plural sample, in which the first constituent appeared only with its plural linking element (*-er-* in the case of *Kind*). The alternate orthography for umlaut vowels (<ae> for ä, <oe> for ö, <ue> for ü) was also included in the samples when searching for the three plural linking elements that involve umlauts.

A foreseeable issue is that the same compound will appear many times in the sample, and even more so if it's highly frequent or even lexicalized (i.e. a fixed form composed of first constituent, possible linking element, and head constituent, e.g. *Mutter-Ø-sprache* 'first language', lit. mother language). Since not all 2000 tokens will be annotated, but rather a randomly selected subset of maximum 100 tokens, it would skew the analysis immensely if many of those 100 tokens were the same highly frequent word. After all, the goal is to analyze the inner construction and internal variability of compound words, which is less likely to be flexible in lexicalized forms.

Therefore, as a measure to negate the influence of lexicalized forms on the analysis, I uniqued the 2000-token samples so that each form only appeared once. This does allow lexicalized compounds to stay in the sample, but this is not a problem for the analysis, since research has shown that German compounds – even lexicalized ones – may be processed as two independent constituents as well as a whole-word form, meaning that internal variation is still possible (Dressler 1987, Libben et al. 2002, Libben 2006:6).

When uniquing samples, one must recognize that the list is not truly "unique" in the sense that each compound appears only once. The same compound could appear multiple times with various number and case inflections on the head, as shown in Table 4.

Table 4: Possible morphological variations in nominal form by declension class

strong nouns	inflection		
nominative/accusative/dative singular	no suffix	*Tag*	'day'
genitive singular	*-s* or *-es* on masculine and neuter nouns	*Tag-es*	'(of the) day'
nominative/accusative/genitive plural	standard plural suffix	*Tag-e*	'days'
dative plural	plural suffix + *-n*, if not already *n*-final	*Tag-e-n*	'days'

weak nouns	inflection		
nominative singular	no suffix	*Held*	'hero'
oblique cases singular, all cases plural	*-en*	*Held-en*	'hero(es)'

This variation is, however, not a problem. In fact, it may be of benefit to the analysis, since now, the samples still reflect word frequency in actual language use, even if only to a small extent, which would otherwise be levelled completely in the uniquing process. Even though a lexicalized compound could hypothetically appear up to four times in a sample (or even five, if both genitive allomorphs are used), this is acceptable, since as I mentioned above, German compounds have been shown to be internally decomposable and the plural effects researched in this study may therefore still be observable. Additionally, since I am not expecting any effect based on nominal case morphology, there is no reason to control for case variation. Therefore, the appearance of multiple inflected forms is an issue to be aware of but it is not a problem for the analysis.

Another consequence of uniquing the samples is a drastic reduction in sample size. After uniquing, samples that were previously 100 to 2000 tokens in size were reduced to anywhere between approximately 25 and 550 tokens. This meant a further elimination of several first constituents from Table 3; this process will be elaborated on in the next section.

2.3 Annotation

The first 100 eligible tokens in each sample were annotated for two binary features: the plurality of the head (plural or singular) and whether the head is a collective noun (yes or no). Not all words in the samples were eligible for annotation because not all of them fit the target description of a first constituent optionally followed by a linking element or subtraction followed by a head constituent. First, I will briefly outline the types of words that were excluded from the annotation, then discuss the consequences for the analysis of the resulting decreased sample size.

Some words were skipped because the corpus' built-in compound analyzer erroneously recognized them as compounds, such as the adjective *fruchtbar* 'fertile', which appeared capitalized in the corpus as *Fruchtbar*. Since all German nouns are written with an initial capital letter, this is a criterion the corpus relies on to identify nouns, which led to the erroneous analysis of the word as *Frucht-Ø-bar* 'fruit bar/pub'.

German also often uses compounds as proper nouns, both as place names (e.g. *Sonne-n-allee* 'sun boulevard': a street in Berlin) and surnames (e.g. *Brett-Ø-meister*, lit. board master). I left these out as well, since proper nouns are internally completely inflexible.

Other compounds were skipped for a morphological reason: their heads do not show plural/singular alternation. This includes most nominalized verbs (e.g. *sein* 'to be' as in *Mensch-Ø-sein* 'personhood'; *werden* 'to become' as in *Frau-Ø-werden* 'becoming a woman'), nouns that have no plural form or technically have one but it is never used (e.g. *Gold* 'gold', **Golde* 'golds'), and pluralia tantum (e.g. *Trümmer* 'ruins').

Occasionally, the plurality of tokens could not be unambiguously determined, even with context, because of German's zero plural suffix and the many syncretisms in adjectival and determiner paradigms. When a token was not clearly singular or plural, it was also left out.

Finally, as I mentioned above, for first constituents that regularly have a figurative as well as a literal interpretation and where occurrence of the figurative meaning is predictable (like *Mutter* 'mother'), the figurative usages were methodically excluded and the literal ones remained.

After these criteria had been applied and the eligible tokens annotated, the sample sizes were reduced even further. To further ensure that the linking element alternation is at least somewhat productive, I introduced another restriction: after uniquing and annotation, each of the two samples for one first constituent had to include at least 30 eligible tokens. This led to the elimination of four first constituents from Table 3: *Stuhl* 'chair', *Baum* 'tree', *Anwalt* 'lawyer', and *Held* 'hero'.

Additionally, in order to keep the scope of this project manageable and to evaluate a similar number of first constituents for each plural linking element, I chose a maximum of six evenly distributed first constituents from those remaining to be part of the analysis. Table 5 shows the final selection of first constituents.

Table 5: Analyzed first constituents sorted by their plural linking element

Umlaut	Umlaut + -e-	Umlaut + -er-	-e-	-er-	-n-	-en-
Mutter 'mother'	Stadt 'city'	Buch 'book'	Hund 'dog'	Kind 'child'	Sonne 'sun'	Frau 'woman'
Vater 'father'	Hand 'hand'	Haus 'house'	Gerät 'device'	Bild 'picture'	Kunde 'customer'	Person 'person'
Apfel 'apple'	Zahn 'tooth'	Bad 'bath(room)'	Weg 'path'	Ei 'egg'	Auge 'eye'	Student 'student'
Nagel 'nail'	Ball 'ball'	Rad 'wheel, bicycle'	Geschenk 'gift'	Lied 'song'	Bauer 'farmer'	Ohr 'ear'
Vogel 'bird'		Schloss 'castle'	Produkt 'product'	Brett 'board'	Katze 'cat'	Dämon 'demon'
		Wurm 'worm'	Brief 'letter'	Schwert 'sword'	Gitarre 'guitar'	Bett 'bed'

These restrictions may seem very complicated, but they are all motivated at the core by finding a balance between two key concerns: ensuring a large enough sample size while still using only those words that show morphological number alternation and whose semantics are predictable based on this alternation. After applying these conditions and criteria, what remains is a group of semantically comparable words that all behave largely in the same way.

Also, since any tendencies and pockets of regularity in linking element distribution are limited to small and very selective groups of words, these tight restrictions help to create a refined data set where patterns would be more observable than among compound words overall.

3 Results

3.1 The effect of the head noun's plural morphology

Two independent binary variables were analyzed in the plural morphology component of this study. The first was morphological number of the head (singular vs. plural) and the second was the type of linking element (non-plural vs. plural). This leads to four possible constructions of compounds:

singular head and non-plural linking element	plural head and non-plural linking element
singular head and plural linking element	plural head and plural linking element

To evaluate the raw data, I used the chi-square test, which shows whether a statistically significant relationship exists between the two independent variables being investigated. If there is a significant relationship between the variables (i.e. the data is sufficiently unlikely to be distributed the way it is due to pure chance), the null hypothesis can be refuted. In this case, the null hypothesis is that no relation exists between the two independent variables, that is, between the head noun's plurality and the choice of linking element.

Initially, I collapsed all the data from the 39 first constituents into a single four-cell table as above. The result of a chi-square test on that data was insignificant ($X^2(1, n = 7145) = 0.008$, $p > 0.05$ ($p = 0.9273$)). The desired significance level was $\alpha < 0.05$, and here, the p-value is 0.9; there was nowhere near a significant relationship between these variables, so the null hypothesis could not be refuted.

In case the merging of all data eliminated individual nuances, I performed chi-square tests for each of the 39 first constituents separately. An example of the resulting four-field table for the

first constituent *Vater* 'father' is given in Table 6 below. (See Appendix A for the raw data and example compounds for all first constituents.)

Table 6: Raw data for *Vater* 'father'

	Compounds with a singular head		Compounds with a plural head	
Compounds with a non-plural linking element	Vater-Ø-rolle 'father role'	(76)	Vater-Ø-gefühl-e 'fatherly feelings'	(24)
Compounds with a plural linking element	Väter-beteiligung 'fathers' participation'	(62)	Väter-recht-e 'fathers' rights'	(38)

In 31 of 39 cases, this null hypothesis could still not be refuted. The 8 first constituents that did reach a level of significance, ordered by p-value, are:

Geschenk 'gift' $X^2(1, n = 200) = 10.569, p < 0.05 (p = 0.001)$
Apfel 'apple' $X^2(1, n = 166) = 8.39, p < 0.05 (p = 0.004)$
Produkt 'product' $X^2(1, n = 200) = 7.587, p < 0.05 (p = 0.006)$
Stadt 'city' $X^2(1, n = 200) = 6.281, p < 0.05 (p = 0.012)$
Bad 'bath(room)' $X^2(1, n = 200) = 5.159, p < 0.05 (p = 0.023)$
Ball 'ball' $X^2(1, n = 135) = 5.028, p < 0.05 (p = 0.025)$
Person 'person' $X^2(1, n = 200) = 4.542, p < 0.05 (p = 0.033)$
Vater 'father' $X^2(1, n = 200) = 3.95, p < 0.05 (p = 0.047)$

However, only four of these eight (*Apfel* 'apple', *Stadt* 'city', *Person* 'person' and *Vater* 'father') demonstrate the desired relationship, where morphological plurality of the head co-occurs with an increased use of the plural linking element over the non-plural one(s). For example, both the compound *Vater-Ø-erfahrung* 'father's experiences' and its normal plural form *Vater-Ø-erfahrung-en* (with no internal changes) appear in the corpus. The form *Väter-erfahrung-en* is also documented, where the head's plural suffix co-occurs with the plural linking

element (in this case, umlaut), though the hypothetical form *Väter-erfahrung*, with a singular

head and a plural linking element, does not exist in the corpus.

The other four first constituents, *Geschenk* 'gift', *Produkt* 'product', *Bad* 'bath(room)',

and *Ball* 'ball' show (sometimes very strongly) the opposite relation, where a plural linking

element more frequently appears in a compound with a morphologically singular head. The most

extreme case is that of *Bad* 'bath(room)', where 82% of the instances of the plural linking

element appear with a singular head.

Table 7: Raw data for *Bad* 'bath(room)'

	Compounds with a singular head		Compounds with a plural head	
Compounds with a non-plural linking element	Bad-Ø-fenster 'bathroom window'	(67)	Bad-Ø-ruine-n 'ruins of baths'	(33)
Compounds with a plural linking element	Bäd-er-skandal 'bath scandal'	(82)	Bäd-er-anlage-n 'bath facilities'	(18)

The Cramér's Phi effect size test was performed on the four first constituents that showed

a significant result and the desired distribution. This test shows to what extent the relationship

between the variables explains the variation in the results. For all four, the effect strength was

small (quantified as $0.10 < \phi < 0.30$ (Gravetter & Wallnau 2007:603)) but present.

Apfel 'apple'	$\phi = 0.225$
Stadt 'city'	$\phi = 0.177$
Person 'person'	$\phi = 0.151$
Vater 'father'	$\phi = 0.141$

This means that the observed relationship explains between 14.1% and 22.5% of the variance in these data sets and the rest of the variance is due to other factors.

In summary, there is overall no significant indication in this data which would suggest that morphologically plural compounds are more likely to contain the appropriate plural linking element for the first constituent.

3.2 The effect of the head noun's plural semantics

The second component of this corpus study tested the co-occurrence of collective head nouns with plural linking elements. Two independent variables were analyzed: whether the head noun was a collective noun or not (regardless of grammatical plurality) and whether the linking element was plural or non-plural. The null hypothesis was that there is no relation between these two variables.

The correlation between these two variables is strong throughout the data. Compound heads that were collective nouns co-occurred significantly more with a plural linking element than with a non-plural one ($X^2(1, n = 7145) = 157.47, p < 2.2 \times 10^{-16}$ ($p = 4.05 \times 10^{-36}$)). This allows us to refute the null hypothesis and state that there is indeed a relationship between semantic plurality of the compound's head and the selection of a plural linking element over a non-plural one. Almost 15% of the variance in the data was explained by this correlation ($\phi = 0.148$) which is a small but notable effect strength.

As illustration, of the 21 instances of the first constituent *Haus* 'house' appearing with a collective head constituent, in 19 of those compounds (90.48%), *Haus* takes its plural linking element umlaut + *-er-*. Forms like *Häus-er-meer* 'sea of houses', *Häus-er-liste* 'list of houses', and *Häus-er-reihe* 'row of houses' are attested, while the corresponding non-plural forms *Haus-Ø-meer, Haus-Ø-liste, Haus-Ø-reihe* are not.

Interestingly, the relationship between collective heads and plural linking elements in the subset of first constituents with human referents (*Mutter* 'mother', *Vater* 'father', *Kind* 'child', *Kunde* 'customer', *Bauer* 'farmer', *Frau* 'woman', *Person* 'person', and *Student* 'student') is even more substantial than in the overall data. It is similarly highly significant $(X^2(1, n = 1529) = 69.1, p < 2.2 \times 10^{-16}$ ($p = 9.36 \times 10^{-17}$)) but the effect is even stronger, explaining 21.3% of the total variance ($\phi = 0.213$). Appendix B contains more example compounds and the raw data from the semantics component of the study.

In summary, there is a significant correlation between the semantic plurality of the compound head and the selection of a plural linking element over a non-plural one.

4 Discussion

This corpus study has allowed me to untangle two potential sources of linking element conditioning, morphology and semantics, in my study of the distribution of plural linking elements. I have shown that the pattern of linking element alternation among concrete count nouns seems to be motivated semantically rather than morphologically, with further semantic differentiation between first constituents with human and non-human referents.

The relationship between semantic plurality of the head and plural linking elements can be characterized as follows: when the first constituent is a concrete count noun and the compound head is a collective noun, it is very likely that the linking element will be homophonous with the first constituent's plural marker (i.e. will be a plural linking element). In 73.79% of all compounds with collective heads in the data (411 of 557 total collective heads), the linking element was a plural one.

The correlation between plural linking elements and collective heads is even higher in the samples where the first constituent has a human referent: 81.67% of these collective-headed compounds include a plural linking element (147 of 180).

However, as the low scores in the Cramér's Phi effect strength test shows us, much of the variation in the data remains unexplained by these relationships, which reflects the large amount of variability and complexity that exists within linking element systems.

In contrast to the semantic conditioning, linking element alternation was shown to be not at all conditioned by plural morphological inflection of the noun's head. Below are several possible reasons why a significant effect was not observable in the overall data.

1. This phenomenon may be more semantically restricted than this data set could test. Based on the previously cited examples of *Frau-en-stimme* 'woman's voice', *Mutter-Ø-herz*

'mother's heart', and *Vater-Ø-erfahrung* 'father's experience', this effect is possibly only apparent in compounds with clearly defined semantics: the first constituent must denote a possessor of exactly one (possibly inalienable) thing and the head noun must denote that thing. A more semantically restricted data set would be required to test this hypothesis.

2. The normal practice of pluralizing compounds by inflecting the head without any internal change (see Section 1.3.1) may be too strong to expect regular, productive compound-internal alternation effects that are conditioned by the inflections of the head.

3. The preference for one linking element over another may be too strong among most of the selected first constituents. Schwa-final first constituents like *Sonne* 'sun', for example, are generally under phonological pressure to close the schwa syllable with the linking element *-n-*, which in most cases is their plural linking element (e.g. *Sonne-n-brille* 'sunglasses'). Any tendency toward a particular linking element makes productive alternation less likely. The criterion designed to ensure reliable alternation (>100 tokens from each sample) was probably overly permissive, and this lower cut-off should be raised if designing a similar experiment in the future. Alternately, the acceptability of alternation could be judged by native speakers. (However, the lack of apparent alternation doesn't necessarily mean that there are no semantic effects from the second constituent's plurality; recall the examples *Frau-en-stimme* 'woman's voice' versus *Frau-en-stimme-n* 'women's voices' from 17. There is a semantic change in the plurality of the first constituent, but it's not reflected in the linking element because of *Frau*'s strong tendency to occur with *-en-*. This type of experiment that is limited to formal alternations cannot detect these changes, so a psycholinguistic experiment would be needed to see what change, if any, speakers perceive between *Frau-en-stimme* and *Frau-en-stimme-n*.)

4. Semantic variability in first constituents that I did not account for, e.g. homophony, may also have been a confounding factor. Different linking elements on homophonic first constituents can sometimes change which of the multiple meanings is brought to bear in that compound (cf. -Ø- after *Maus* creates the meaning of a computer mouse, as in *Maus-Ø-klick* 'mouse click' or *Maus-Ø-taste* 'mouse button', while umlaut + *-e-* is used when denoting the rodent, as in *Mäus-e-baby* 'baby mouse' or *Mäus-e-falle* 'mouse trap'). None of the cases of homophony in this data were as clear-cut as with *Maus*, but the various meanings that are conveyed by the same form may nonetheless have confounded the analysis. Future experiments with this design should exclude homophonic first constituents.

For the rest of this section, I will discuss possible reasons why four first constituents emerged from the morphological conditioning section of the study with a statistically significant distribution in the opposite direction than expected.

Two of these first constituents, *Geschenk* 'gift' and *Produkt* 'product', also showed no strong alternation effect in the semantically-conditioned part of the experiment. The average co-occurrence percentage of plural linking elements and collective heads is approximately 74%; *Geschenk*'s percentage was much lower at 60% (6 of 10), and *Produkt*'s was lower still, only 52.38% (11 of 21). This suggests that, while these two first constituents do occur both with their plural and non-plural linking elements, the alternation between them is not related to semantic plurality.

Bad 'bath(room)' and *Ball* 'ball' do show the plural linking element/collective head co-occurrence effect (*Bad* with 80%, 8 of 10; *Ball* with 84%, 11 of 13), but their homophony could help explain why they behaved differently than the other nouns in the sample, since they both

representing multiple meanings. *Bad* means 'bath' as well as 'bathroom', and *Ball* refers to 'ball' in two of the senses that it also has in English: any spherical object, as well as the formal social event. These multiple meanings may have confounded the analysis.

In the following sections, I will relate the findings of this study to the three positions in the literature outlined in section 1.2.1 above, comment on what these conclusions may mean for our conception of the internal structure of German compound words, and finally suggest directions for future research in this fascinating and still under-investigated area.

4.1 Response to the literature

Position 1: No relation exists between any linking elements and plural semantics.

It is true that plural semantics are not universally associated with plural linking elements, but this research has revealed an area of fairly regular co-occurrence: when the meaning of the second constituent forces the first constituent to take on plural meaning. Thus, this study has shown that linking elements are indeed correlated with plural semantics in a compound, which speaks against this analysis of across-the-board semantic emptiness.

However, the caution shown in this approach is commendable; since not all plural linking elements co-occur with plural meaning, some distinction between linking elements with plural form and meaning and those with only plural form must be made.

Position 2: Some linking elements carry and convey plural meaning.

This study corroborates Position 2 by showing that some linking elements (more specifically, some plural linking elements) carry and convey plural meaning, because they are more likely to appear in situations of plural semantics than non-plural linking elements are. These results help to

fill in part of the primary gap in this theory, which is the question of when linking elements express plural meaning and when they do not.

Position 3: When a plural linking element co-occurs with plural meaning of the first constituent, it should actually be considered a plural inflectional morpheme that can appear in compounds. This conceptual distinction between inflectional and derivational morphology within compounds is more theoretical than can be targeted with any certainty by this study. It's also impossible to say with any certainty whether Dressler's (1987) concept of "word-form-based compounds" (1987:75) may be true.

However, I don't find this suggestion unrealistic. The concept of plural morphemes existing within compounds was, after all, only disproven in the oral modality by Koester and colleagues (2004), so it's possible that these plural linking elements may have quasi-morphemic status; I will discuss this idea further in the next section.

4.2 Internal inflection and German compounding morphology

This analysis supports a distinction between the plural linking elements that co-occur with plural meaning and the plural and non-plural linking elements that do not (i.e. that are semantically empty, at least regarding number). The strong co-occurrence of plural form and plural meaning and its significantly predictable distribution (at least within this data set) speak for the third analysis, that there may be inflection occurring within German nominal compounds.

I tend to a more moderate position than Dressler (1987) and Fehringer (2009), however, since the co-occurrence of plural form and meaning does not happen in every affected compound (cf. *Bild-Ø-sammlung* vs. *Bild-er-sammlung* 'picture collection', which are both documented). This non-universality is characteristic of the irregularity and lack of definite rules of the linking element system and means that we should exercise caution in claiming that all plural linking

elements that co-occur with plural semantics are plural morphemes, since if they were, their presence would be entirely predictable. Therefore, I'd like to suggest that the linking elements that we see in compounds like *Bild-er-sammlung* 'picture collection' and *Frau-en-mannschaft* 'women's team' are probably not plural inflectional morphemes, but rather, plural-morpheme-like linking elements. These are prompted by the plural semantics forced upon the first constituent by the head and possibly produced through analogy to non-plural linking elements that share the same form.

As additional evidence, Fehringer (2009) observes that the increased occurrence of the plural -*s* in colloquial northern German dialects leads to a more frequent usage of the linking element -*s*- in compounds with collective heads like *Mädel-s-party* 'girls' party'. This correlation suggests a close relationship between plural-morpheme-like linking elements and normal plural morphemes.

4.3 Directions for future research

4.3.1 Further research using existing data

Since semantics has been shown to be a good indicator of linking element distribution, it would be interesting to explore other semantic sub-groupings within the existing data. For example, nouns denoting objects that occur primarily in pairs or groups (for example, body parts like hands, teeth, and eyes) often behave differently in constructions involving plurality than other nouns (see Nenonen & Niemi 2010) and could show other interesting effects.

4.3.2 On the relationship of plurality and linking elements

Both Positions 2 and 3 separate linking elements into the same two basic categories: those that convey plural meaning and those that do not. However, Position 2 does it with less specific

and descriptive terminology and may thus overlook the not-unprecedented possibility of compound-internal inflection. Even if one prefers the more intuitively straightforward analysis of Position 2, that all interfixes count as linking elements and that some may express plural meaning, a language user likely cannot arrive at these plurality-expressing linking elements that have the same form as plural suffixes without mentally accessing and activating the plural suffixes themselves. A psycholinguistic experiment to study the extent of plural morpheme activation in the written modality in such compounds would be highly informative in determining the extent to which these linking elements really are like plural morphemes. This would be an excellent parallel to the investigation of aural perception of plural linking elements by Koester and colleagues (2004).

To my knowledge, there are few other studies that investigate linking element alternation, with or without a focus on plurality. Further research on productive linking element alternation in minimal pairs, like those in 14 and 15, would be fascinating. Nübling and Szczepaniak (2013) and Fuhrhop (1996) have suggested certain types of words that allow productive plural/non-plural linking element alternation (see Section 1.3). However, in the present study, many first constituents that do not match the proposed descriptions still show an alternation effect, which suggests that more first constituents may allow productive alternation than previously suggested. To study this, one could apply the same plural/non-plural alternation to first constituents with varying phonological and morphological characteristics and gather native speaker judgments, in order to discover exactly which first constituents allow alternation and which do not. Such an investigation would complement the current study, which looked at alternation being caused by characteristics of the head, by investigating speaker- and context-based linking element alternation coming from the first constituent. It would also be extremely helpful for the study of

linking element distribution in general, because it would clarify what factors and characteristics of the first constituent allow certain linking elements while disallowing others.

Finally, to test if the morphologically conditioned plurality effect does exist, a study similar in structure to the current one should be performed with a more semantically restricted data set.

5 Conclusion

This study has shown one more small area of relative regularity in the complex distribution of linking elements in German nominal compounds. Linking elements that are homophonic with the first constituent's plural suffix occur in 74% of compounds where the head is a collective noun, i.e. where the referent of the first constituent is semantically plural. This shows that certain linking elements do co-occur with plural meaning, suggesting that a distinction among linking elements between those that are semantically empty (e.g. *-n-* in *Sonne-n-brille* 'sunglasses') and those with plural form that co-occur with plural meaning (e.g. *-er-* in *Bild-er-sammlung* 'picture collection') should be made. I propose that, in cases of co-occurrence of plural form and meaning, linking elements should be considered plural-morpheme-like. Further investigation is required to determine the extent to which these linking elements can be considered plural morphemes in the written modality.

Researchers of linking elements who work only with generalizations and overarching statements (e.g. Neef (2015), among others) often see counterexamples and distributions that are not utterly predictable as reasons to disregard any small patterns that do exist in the linking element system. However, it's necessary to keep in mind that the distributional system does have tendencies and small areas of regularity, which are important to understand in order to expand our knowledge of German compounding.

REFERENCES

DRESSLER, WOLFGANG U. 1987. Morphological islands: constraints or preference? *Language Topics*, ed. by R. Steele & T. Treadgold, 71–79. Amsterdam: Benjamins.

FEHRINGER, CAROL. 2009. Wie wär's mit einem richtigen Mädelsabend? Plural -s within Compounds in Colloquial Northern German. *Journal of Germanic Linguistics* 21.2. 149–165.

FUHRHOP, NANNA. 1996. Fugenelemente. *Deutsch - Typologisch,* 525–549. Berlin: Walter De Gruyter.

GALLMANN, PETER. 1999. Fugenmorpheme Als Nicht-Kasus-Suffixe. *Variation Und Stabilitat in Der Wortstruktur: Untersuchungen Zu Entwicklung, Erwerb Und Varietaten Des Deutschen Und Anderer Sprachen,* ed. by Matthias Butt and Nanna Fuhrhop, 177–190. Hildesheim: G. Olms Verlag.

GRAVETTER, FREDERICK, & LARRY WALLNAU. 2007. *Statistics for the Behavioral Sciences.* 7th edn. Toronto: Thomson Wadsworth.

KOESTER, DIRK, THOMAS C. GUNTER, SUSANNE WAGNER, & ANGELA D. FRIEDERICI. 2004. Morphosyntax, Prosody, and Linking Elements: The Auditory Processing of German Nominal Compounds. *Journal of Cognitive Neuroscience* 16.9. 1647–1668.

KROTT, ANDREA, ROBERT SCHREUDER, R. HARALD BAAYEN, & WOLFGANG U. DRESSLER. 2007. Analogical Effects on Linking Elements in German Compound Words. *Language and Cognitive Processes* 22.1. 25–57.

LIBBEN, GARY. 2006. Why Study Compound Processing? An overview of the issues. *The Representation and Processing of Compound Words*, ed. by Gary Libben and Gonia Jarema, 1–22. Oxford University Press.

LIBBEN, GARY, GONIA JAREMA, WOLFGANG U. DRESSLER, JACQUELINE STARK, & CHRISTIANE
PONS. 2002. Triangulating the Effects of Interfixation in the Processing of German
Compounds. *Folia Linguistica* 36.1-2. 23–43.

NEEF, MARTIN. 2015. The Status of So-called Linking Elements in German: Arguments in Favor
of a Non-functional Analysis. *Word Structure* 8.1. 29–52.

NENONEN, MARJA, & JUSSI NIEMI. 2010. Mismatches between grammatical number and
conceptual numerosity: A number-decision experiment on collective nouns, number
neutralization, pluralia tantum, and idiomatic plurals. *Folia Linguistica* 44.1. 103–125.

NÜBLING, DAMARIS, & RENATA SZCZEPANIAK. 2013. Linking Elements in German Origin,
Change, Functionalization. *Morphology* 23.1. 67–89.

SCHÄFER, ROLAND, & FELIX BILDHAUER. 2012. Building Large Corpora from the Web Using a
New Efficient Tool Chain. *Proceedings of the Eighth International Conference on
Language Resources and Evaluation*. 486–493.

WEGENER, HEIDE. 2003. Entwicklung und Funktion der Fugenelemente im Deutschen. Oder:
Warum wir keine *Autosbahn haben. *Linguistische Berichte* 196. 425–456.

Morphological plurality raw data and example compounds from DECOW16A.
Not all cited examples may be acceptable to all native German speakers. Also, note that the English translations are idiomatic rather than literal and thus don't necessarily express the plurality of the first constituent as the German compound does. Finally, in all cases where the plural zero suffix is used, it was clear from the context that the compound is grammatically plural.

i. Plural linking element (LE): Umlaut

Mutter 'mother'	*Singular Head*		*Plural Head*	
Non-Plural LE	Mutter-Ø-liebe 'mother's love'	(76)	Mutter-Ø-tier-e 'mother animals'	(24)
Plural LE	Mütter-zentrum 'mother's centre'	(84)	Mütter-seminar-e 'seminars for mothers'	(16)

Vater 'father'	*Singular Head*		*Plural Head*	
Non-Plural LE	Vater-Ø-rolle 'father role'	(76)	Vater-Ø-gefühl-e 'fatherly feelings'	(24)
Plural LE	Väter-beteiligung 'fathers' participation'	(62)	Väter-recht-e 'fathers' rights'	(38)

Apfel 'apple'	*Singular Head*		*Plural Head*	
Non-Plural LE	Apfel-Ø-tag 'apple day'	(65)	Apfel-Ø-scheibe-n 'apple slices'	(35)
Plural LE	Äpfel-diebstahl 'apple theft'	(27)	Äpfel-sämling-e 'apple seedlings'	(39)

Nagel 'nail'	*Singular Head*		*Plural Head*	
Non-Plural LE	Nagel-Ø-lack 'nail polish'	(56)	Nagel-Ø-ränd-er 'nail edges'	(44)
Plural LE	Nägel-größe 'nail size'	(18)	Nägel-abdrück-e 'nail prints'	(29)

Vogel 'bird'	*Singular Head*		*Plural Head*	
Non-Plural LE	Vogel-Ø-perspektive 'bird's eye view'	(63)	Vogel-Ø-foto-s 'photos of birds'	(37)
Plural LE	Vögel-käfig 'birdcage'	(22)	Vögel-nest-er 'bird's nests'	(25)

ii. Plural linking element: Umlaut + *-e-*

Stadt 'city'	Singular Head		Plural Head	
Non-Plural LE	Stadt-Ø-zentrum 'city centre'	(80)	Stadt-Ø-mauer-n 'city walls'	(20)
Plural LE	Städt-e-meisterschaft 'city championship'	(63)	Städt-e-zerstörer-Ø 'destroyers of cities'	(37)

Hand 'hand'	Singular Head		Plural Head	
Non-Plural LE	Hand-Ø-fläche 'palm of the hand'	(68)	Hand-Ø-bewegung-en 'hand movements'	(32)
Plural LE	Händ-e-druck 'handshake'	(49)	Händ-e-paar-e 'pairs of hands'	(20)

Zahn 'tooth'	Singular Head		Plural Head	
Non-Plural LE	Zahn-Ø-spange 'braces'	(71)	Zahn-Ø-problem-e 'tooth problems'	(36)
Plural LE	Zähn-e-behandlung 'tooth treatment'	(25)	Zähn-e-spezialist-en 'tooth specialists'	(9)

Ball 'ball'	Singular Head		Plural Head	
Non-Plural LE	Ball-Ø-spiel 'game involving a ball'	(67)	Ball-Ø-verlust-e 'loss of ball possession'	(33)
Plural LE	Bäll-e-dieb 'ball thief'	(31)	Bäll-e-bäd-er 'ball pits'	(4)

iii. Plural linking element: Umlaut + *-er-*

Buch 'book'	*Singular Head*		*Plural Head*	
Non-Plural LE	Buch-Ø-laden 'bookstore'	(81)	Buch-Ø-verlag-e 'book publishers'	(19)
Plural LE	Büch-er-regal 'bookshelf'	(73)	Büch-er-würm-er 'bookworms'	(27)

Haus 'house'	*Singular Head*		*Plural Head*	
Non-Plural LE	Haus-Ø-hälfte 'half of a house'	(59)	Haus-Ø-schuh-e 'slippers, house shoes'	(41)
Plural LE	Häus-er-labyrinth 'labyrinth of houses'	(46)	Häus-er-preis-e 'house prices'	(54)

Bad 'bath-(room)'	*Singular Head*		*Plural Head*	
Non-Plural LE	Bad-Ø-fenster 'bathroom window'	(67)	Bad-Ø-ruine-n 'ruins of baths'	(33)
Plural LE	Bäd-er-skandal 'bath scandal'	(82)	Bäd-er-anlage-n 'bath facilities'	(18)

Rad 'wheel, bicycle'	*Singular Head*		*Plural Head*	
Non-Plural LE	Rad-Ø-weg 'bicycle path'	(59)	Rad-Ø-fahrzeug-e 'wheeled vehicles'	(41)
Plural LE	Räd-er-werk 'wheelwork'	(71)	Räd-er-fläche-n 'surfaces of wheels'	(29)

Schloss 'castle'	*Singular Head*		*Plural Head*	
Non-Plural LE	Schloss-Ø-herr 'lord of the castle'	(85)	Schloss-Ø-türm-e 'castle towers'	(15)
Plural LE	Schlöss-er-gelände 'castle grounds'	(86)	Schlöss-er-reise-n 'vacations to castles'	(14)

Wurm 'worm'	*Singular Head*		*Plural Head*	
Non-Plural LE	Wurm-Ø-loch 'wormhole'	(62)	Wurm-Ø-larv-en 'worm larvae'	(38)
Plural LE	Würm-er-gott 'worm god'	(30)	Würm-er-art-en 'worm species'	(17)

iv. Plural linking element: *-e-*

Hund 'dog'	*Singular Head*		*Plural Head*	
Non-Plural LE	Hund-s-kopf 'dog's head'	(80)	Hund-Ø-haar-e 'dog hair'	(20)
Plural LE	Hund-e-rasse 'dog species'	(68)	Hund-e-besitzer-Ø 'dog owners'	(32)

Gerät 'device'	*Singular Head*		*Plural Head*	
Non-Plural LE	Gerät-Ø-größe 'device size'	(65)	Gerät-Ø-einstellung-en 'device settings'	(34)
Plural LE	Gerät-e-rückseite 'device backside'	(68)	Gerät-e-kapazität-en 'device capacity'	(32)

Weg 'path'	*Singular Head*		*Plural Head*	
Non-Plural LE	Weg-Ø-marker 'path marker'	(61)	Weg-Ø-ränd-er 'edges of the path'	(39)
Plural LE	Weg-e-recht 'right of way'	(64)	Weg-e-information-en 'path information'	(36)

Geschenk 'gift'	*Singular Head*		*Plural Head*	
Non-Plural LE	Geschenk-Ø-karton 'gift box'	(53)	Geschenk-Ø-idee-n 'gift ideas'	(47)
Plural LE	Geschenk-e-sack 'bag of gifts'	(76)	Geschenk-e-tipp-s 'tips about gifts'	(24)

Produkt 'product'	*Singular Head*		*Plural Head*	
Non-Plural LE	Produkt-Ø-angebot 'product offer'	(51)	Produkt-Ø-detail-s 'product details'	(49)
Plural LE	Produkt-e-temperatur 'product temperature'	(71)	Produkt-e-linie-n 'product lines'	(29)

Brief 'letter'	*Singular Head*		*Plural Head*	
Non-Plural LE	Brief-Ø-papier 'letter paper'	(71)	Brief-Ø-umschläg-e 'envelopes'	(29)
Plural LE	Brief-e-schreiberin 'female letter writer'	(25)	Brief-e-leser-Ø 'letter readers'	(11)

v. Plural linking element: *-er-*

Kind 'child'	*Singular Head*		*Plural Head*	
Non-Plural LE	Kind-s-braut 'child bride'	(70)	Kind-s-tötung-en 'child murders'	(30)
Plural LE	Kind-er-arzt 'pediatrician'	(62)	Kind-er-schuh-e 'children's shoes'	(38)

Bild 'picture'	*Singular Head*		*Plural Head*	
Non-Plural LE	Bild-Ø-interpretation 'picture interpretation'	(71)	Bild-Ø-entwürf-e 'drafts of a picture'	(29)
Plural LE	Bild-er-saal 'hall of pictures'	(65)	Bild-er-dieb-e 'picture thieves'	(35)

Ei 'egg'	*Singular Head*		*Plural Head*	
Non-Plural LE	Ei-Ø-gelb 'egg yolk'	(54)	Ei-Ø-leiter-Ø 'Fallopian tubes'	(46)
Plural LE	Ei-er-farbe 'egg colour'	(57)	Ei-er-karton-s 'egg cartons'	(43)

Lied 'song'	*Singular Head*		*Plural Head*	
Non-Plural LE	Lied-Ø-text 'song text'	(58)	Lied-Ø-zeile-n 'lines of a song'	(42)
Plural LE	Lied-er-komponist 'composer'	(66)	Lied-er-titel-Ø 'song titles'	(34)

Brett 'board'	*Singular Head*		*Plural Head*	
Non-Plural LE	Brett-Ø-spiel 'board game'	(68)	Brett-Ø-schicht-en 'layers of boards'	(32)
Plural LE	Brett-er-kreuz 'cross made of boards'	(63)	Brett-er-zäun-e 'board fence'	(37)

Schwert 'sword'	*Singular Head*		*Plural Head*	
Non-Plural LE	Schwert-Ø-knauf 'pommel'	(65)	Schwert-Ø-scheide-n 'scabbards'	(35)
Plural LE	Schwert-er-schlag 'sword hit'	(44)	Schwert-er-spitze-n 'sword point'	(18)

vi. Plural linking element: *-n-*

Sonne 'sun'	Singular Head		Plural Head	
Non-Plural LE	Sonne-Ø-scheibe 'solar disc'	(69)	Sonne-Ø-blume-n 'sunflowers'	(31)
Plural LE	Sonne-n-aufgang 'sunrise'	(73)	Sonne-n-schirm-e 'sunshades'	(27)

Kunde 'customer'	Singular Head		Plural Head	
Non-Plural LE	Kunde-Ø-beratung 'customer consulting'	(66)	Kunde-Ø-meinung-en 'customers' opinions'	(34)
Plural LE	Kunde-n-befragung 'customer survey'	(61)	Kunde-n-problem-e 'customer problems'	(39)

Auge 'eye'	Singular Head		Plural Head	
Non-Plural LE	Aug–apfel 'eyeball'	(56)	Auge-Ø-farbe-n 'eye colours'	(35)
Plural LE	Auge-n-untersuchung 'eye examination'	(48)	Auge-n-braue-n 'eyebrows'	(52)

Bauer 'farmer'	Singular Head		Plural Head	
Non-Plural LE	Bauer-s-frau 'farmer's wife'	(68)	Bauer-Ø-hochzeit 'farmer marriage'	(32)
Plural LE	Bauer-n-hof 'farm'	(73)	Bauer-n-regel-n 'farmers' rules'	(27)

Katze 'cat'	Singular Head		Plural Head	
Non-Plural LE	Katze-Ø-klo 'cat toilet'	(45)	Katze-Ø-ohr-en 'cat ears'	(11)
Plural LE	Katze-n-liebhaber 'cat lover'	(66)	Katze-n-bild-er 'cat pictures'	(34)

Gitarre 'guitar'	Singular Head		Plural Head	
Non-Plural LE	Gitarre-Ø-studium 'guitar program'	(43)	Gitarre-Ø-töne-n 'guitar sounds'	(30)
Plural LE	Gitarre-n-schule 'guitar school'	(56)	Gitarre-n-riff-s 'guitar riffs'	(44)

vii. Plural linking element: *-en-*

Frau 'woman'	*Singular Head*		*Plural Head*	
Non-Plural LE	Frau-Ø-gefängnis 'women's prison'	(50)	Frau-Ø-gründ-e 'woman reasons'	(21)
Plural LE	Frau-en-körper 'female body'	(73)	Frau-en-klamotte-n 'women's clothes'	(27)

Person 'person'	*Singular Head*		*Plural Head*	
Non-Plural LE	Person-Ø-nummer 'personal number'	(76)	Person-Ø-merkmal-e 'person's characteristics'	(24)
Plural LE	Person-en-zug 'passenger train'	(61)	Person-en-dat-en 'personal data'	(39)

Student 'student'	*Singular Head*		*Plural Head*	
Non-Plural LE	Student-Ø-zimmer 'student room'	(43)	Student-Ø-job-s 'student jobs'	(15)
Plural LE	Student-en-rebellion 'student rebellion'	(74)	Student-en-städt-e 'student cities'	(26)

Ohr 'ear'	*Singular Head*		*Plural Head*	
Non-Plural LE	Ohr-Ø-piercing 'ear piercing'	(53)	Ohr-Ø-problem-e 'ear problems'	(47)
Plural LE	Ohr-en-zeugin 'female earwitness'	(57)	Ohr-en-stöpsel-Ø 'earplugs'	(43)

Dämon 'demon'	*Singular Head*		*Plural Head*	
Non-Plural LE	Dämon-Ø-modus 'demon mode'	(31)	Dämon-Ø-statue-n 'demon statues'	(16)
Plural LE	Dämon-en-peitsche 'demon whip'	(73)	Dämon-en-hund-e 'demon dogs'	(27)

Bett 'bed'	*Singular Head*		*Plural Head*	
Non-Plural LE	Bett-Ø-rand 'edge of the bed'	(64)	Bett-Ø-wanze-n 'bedbugs'	(36)
Plural LE	Bett-en-museum 'bed museum'	(76)	Bett-en-modell-e 'bed models'	(24)

APPENDIX B

Semantic plurality raw data and example compounds from DECOW16A.

All data	*Non-Collective Head*		*Collective Head*	
Non-Plural LE	Haus-Ø-fassade 'house façade'	(3556)	Bild-Ø-vergleich 'comparison of pictures'	(146)
Plural LE	Bauer-n-junge 'farmer boy'	(3032)	Lied-er-sammlung 'song collection'	(411)

Human referents	*Non-Collective Head*		*Collective Head*	
Non-Plural LE	Vater-Ø-komplex 'father complex'	(696)	Kind-Ø-gruppe 'child group'	(33)
Plural LE	Kunde-n-konto 'customer account'	(653)	Frau-en-mannschaft 'women's team'	(147)

Examples of true collectives: *Ansammlung* 'gathering, cluster', *Bund* 'association', *Gruppe* 'group', *Häufung* 'accumulation', *Liste* 'list', *Meeting* 'meeting', *Organisation* 'organization', *Reihe* 'row', *Sammlung* 'collection', *Treffen* 'meeting', *Verband* 'association, union', *Verein* 'society, club'.

Examples of plural action collectives: *Allianz* 'alliance', *Konflikt* 'conflict', *Vergleich* 'comparison', *Wettkampf* 'competition'.