The Role of Information Structure in Pronoun Resolution of Child, Adult L1 and
Adult L2 Speakers

by

Regina Hert

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

Doctor of Philosophy

Department of Linguistics
University of Alberta

Abstract

Research on pronoun resolution has a long-standing tradition and covers a wide range of factors that have been shown to affect the referential link between a pronoun and its referent. The present dissertation focuses on information structural effects in subject and object pronoun resolution in German child and adult speakers as well as second language learners. By doing so, it aims to address to what extent cues of information structure, such as focus marking, can influence pronoun processing and whether the effect differs between the different groups. The dissertation comprises three studies and uses eye-tracking as well as behavioral data. The first study includes child and adult speakers. It investigates the effect of focus marking via *it*-clefts for the processing of referents as well as object pronouns and reflexives. This study also contributes to the clarification of what the underlying issue in children's resolution of object pronouns is. The results suggest that, unlike adults, children's discourse processing is more local which – if the pronoun is not restricted in its referential function – can lead to children's selecting the incorrect referent. The second study manipulates information structure using prosodic focus marking and word order to investigate adults' online use of focus cues in contrast to grammatical role for the resolution of subject and demonstrative pronouns. The results show the personal pronoun to be affected more by subjecthood than focus marking, whereas the demonstrative pronoun is influenced by focus marking to a greater extent than by objecthood, supporting the form-specific approach (Kaiser & Trueswell, 2008). It further addresses the question whether 'prominence' can explain pronoun interpretation preferences. It has been argued that the reason why certain referents are linked with a pronoun is that they are more prominent in the discourse than others (e.g. Arnold, Brown-Schmidt & Trueswell, 2007; von

Heusinger & Schumacher, 2019). In contrast with this assumption, the study suggests that prominence cannot explain why the subject referent is preferred even when the object referent receives more attention than the subject referent. The third study examines how second language speakers of German use information structure for the interpretation of subject pronouns. The results suggest that focus marking in combination with grammatical role leads to a preference emerging in L2 speakers that is similar to L1 speakers found in study 2. Additionally, it tests whether accenting subject and object pronouns affects referent selection in L1 and L2 speakers. L2 speakers seem to rely more on subjecthood during resolution of both pronouns which is similar to L1 speakers with the exception of the accented subject pronoun. Here, L1 speakers select the object referent more often than the subject referent.

Together, these studies give new insights to referent processing in child and adult L1 speakers as well as adult L2 speakers. That is, attention shifts to the focused referent during processing, but that effect does not always proceed to the final interpretation of the pronoun. Precisely this finding is why I argue against prominence driving pronoun resolution, but rather assume a special status for the subject role (cf. Mecklinger, Schriefers, Steinhauer & Friederici, 1995; Meng & Bader, 2000a,b; Schriefers, Friederici & Kühn, 1995). Crucially, the studies show that information structure not only influences referential function of various pronoun types to different degrees but they also show that the extent of the effect of information structure differs in the various participant groups.

Preface

Examining Committee:

Anja Arnhold, Supervisor

Juhani Järvikivi, Supervisor

Jacolien van Rij, Supervisory Committee

Evangelia Daskalaki, Examiner

Elsi Kaiser, External Examiner

The three chapters are intended to be published as separate research articles. Anja Arnhold and Juhani Järvikivi were the supervisory authors and provided conceptual and editorial feedback throughout the writing process.

Chapter 2 has been submitted as: Hert, R., Arnhold, A., & Järvikivi J. (2023). Focus Effect Unveils Children's Local Processing of Pronouns and Reflexives.

Chapter 3 has been submitted as: Hert, R., Järvikivi J., & Arnhold, A. (2023). Nothing else matters – *He* likes subject referents.

Acknowledgements

Contents

List of Tables

List of Figures

Chapter 1

Introduction

With an abundance of pronoun research leading to a variety of outcomes – some contradicting, some supporting a common idea – the present dissertation aims to bring more clarity on factors influencing pronoun resolution. Specifically, the focus of the dissertation is on the effect of information structure. Information structure describes how information is presented within a given discourse (Chafe, 1976). While information structure has been used in pronoun processing before, previous research leaves some questions open that will be described in more detail below and that this dissertation intends to address as well. The way that information structure interacts with pronoun processing seems not to be very clear and results can vary across different studies: within the same language (Järvikivi, van Gompel, Hyönä & Bertram, 2005; Kaiser & Trueswell, 2008), across languages (Colonna, Schimke & Hemforth, 2012, 2014), but also between children and adults (Järvikivi, Pyykkönen-Klauck, Schimke, Colonna & Hemforth, 2014), as well as for first- and second language speakers (Ellert, Roberts & Järvikivi, 2011; Patterson, Esaulova & Felser, 2017).

The use of a pronoun in itself is a cue of information structure, as it signals that its referent is already known or recoverable from the context. However, since pronouns do not encode sufficient information, linking the pronoun to a referent in the discourse enables to retrieve more information about the referent which is needed for the comprehension process. But what happens when there is more than one referent in memory? How does the comprehender choose the correct referent? Previous findings suggest that pronoun resolution can be influenced by a variety

of factors. If information structure can aid in understanding how to link new to old information, can it also influence which referent comprehenders choose for pronouns?

The present dissertation examines the use of information structure in the processing of pronouns and their possible referents in German. In three studies, I investigate the effects of various types of information structural manipulations for the resolution of different pronoun types as well as different participant groups. The purpose of this investigation is to tell us more about the use of information structure during the processing of pronouns and their possible referents and to which degree it may affect the different pronoun types and the different participant groups. In the remainder of the introduction, I first present a brief overview on information structure, followed by an overview of pronoun resolution within the three different participant groups under investigation. I then turn to the goals of the present dissertation, describing its layout in more detail.

1.1   Information Structure: an Overview

According to Chafe (1976), the primary idea of information structure is how information in the discourse context is presented. It can be seen as the packaging of the information. The same content can be packaged in different ways, and it does not directly affect the information content. The assumption is that the speaker adjusts the utterance so that it meets the assumed needs of the addressee. In this sense, information structure reflects what the speaker hypothesizes about the addressee's beliefs (Prince, 1981). In other words, the term information structure stands for the way certain properties of natural language aid the speaker during the communication process to take the addressee's current information state into account in order to ease communication flow (Krifka & Musan, 2013). Information that is mutually shared or assumed to be mutually shared, the common ground, is subject to steady change. The common ground is the basis for how the change unfolds (Krifka & Musan, 2013). Since communication is thought to work similarly in all languages, it is assumed that all languages are equipped with devices that help express infor-

mation structural categories such as focus, topic, newness, and givenness. These devices may be in the form of word order, specific syntactic constructions, specific lexical items, certain particles, and prosodic means. Hereinafter, I will give a short overview of the definition of the different notions that are used throughout this dissertation.

For focus, there are many different ways this concept has been defined, however, one common characteristic is the sense of newness. For instance, Lambrecht (1994) views focus as producing an update in the common ground, and it is contrasted to the notion of presupposition (Chomsky, 1969). In contrast, Dik, Hoffman, de Jong, Sie, Stroomer & de Vries's (1980) definition of focus does not imply that the focused information needs to be new to the discourse, but rather focus marks the relatively most important information. Focus also has been identified as the part of a sentence that answers a preceding (explicit or implicit) *wh*-question (Beaver, Roberts, Simons & Tonhauser, 2017; Büring, 2003; Roberts, 2012). A more general definition has been put forward by Rooth (1985, 1992) who describes focus as the indication of present alternatives which are relevant for the interpretation of linguistic expressions. Note that the presented definitions of focus are not mutually exclusive.

(1)   A: What did Monica lend Rachel?
       B: Monica lent Rachel earrings.

For (1), Dik et al.' (1980) observation would mean that *earrings* would receive focus and therefore be the most important information in the utterance. Following the idea of common ground update (Lambrecht, 1994) and the contrast of the presupposition — Monica lent something to Rachel — (Chomsky, 1969), *earrings* would be new information to the listener. According to the definition of Rooth (1985, 1992), *earrings* would be part of a set of alternatives. For instance, the present alternatives for *what* could be *earrings* and *watch*.

Similar to focus, there are different attempts of defining the notion of topic. It has been described in terms of old or given information of an utterance (Chafe, 1976). The topic constituent determines the entity about which information is

given and under which it should be stored (Krifka, 2008). Often these two defining properties are used synonymously, however Reinhart (1981) sees the old-information property not as a necessary condition for topichood. One straightforward example she gives is the sentence *Felix praised himself* in the context of (2) (Reinhart, 1981, 72):

(2)   A: Who did Felix praise?
         B: Felix praised HIMSELF.

The question would identify *Felix* in B as the topic, since *Felix* has been mentioned in the question, and would be therefore considered "old" information. The answer to the question, *himself*, would be identified as the expression in focus. Assuming that focus is defined as "new" and topic as "old" information, this would contradictorily make the the referent of *Felix* both old and new information. According to Reinhart (1981), new information is not just added to the common ground content in the form of unstructured propositions, but it is rather associated with entities. Following her file card metaphor, topics can be seen as the title of a file card under which the new information is stored.

Topic continuity – the topic of one sentence is the same as the topic of the preceding sentence – is motivated by a higher predictability for continuous elements (topics) and higher predictability results in easier processing (Givón, 1983). Givón (1983) proposed a hierarchy of referential expression which encodes the degree of topic continuity. Other theories, such as *Accessibility Theory* (Ariel, 1988, 1990, 2001) or the *Givenness Hierarchy* (Gundel, Hedberg & Zacharski, 1993), have proposed similar ideas regarding the accessibility or cognitive status encoded in referential expression. The common idea is that the different referential expressions are chosen according to the current recoverability of the intended referent in the addressee's mind (Gundel, 2003). The less recoverable a referent is, the more explicit information is available for its recoverability (Givón, 1983). In example (3a), the use of the full noun phrase *the constructors* implies that the referent is not easily recoverable from the previous discourse and therefore is not easily accessible in memory, for instance, either because the referent has not been mentioned in the discourse yet or that the referent has been mentioned, but other potential referents

would interfere. In example (3b), the use of the pronoun *they* indicates that its referent is easily recoverable from the preceding discourse, e.g., *The constructors have started to work on the first floor*.

(3)   I woke up early, because

     a.   the constructors were making too much noise.

     b.   they were making too much noise.

The questions that emerge are how does the listener/reader decide on a referent? Particularly, what makes one referent more easily recoverable than another? In other words, what makes one referent more likely to be linked to a pronoun than another referent? The next section gives an overview of studies addressing the role of information structure in the interpretation of subject and object pronouns.

1.2   Pronoun Resolution and Its Interplay with Information Structure

While research on pronoun resolution has included a wide range of factors that may impact the referential link between a pronoun and its referent, in this dissertation the focus is on grammatical role (e.g. Crawley & Stevenson, 1990; Frederiksen, 1981; Fukumura & van Gompel, 2015; Gordon & Chan, 1995; Kaiser, 2011a; Song & Fisher, 2005), order of mention (e.g. Gernsbacher, 1985; Gernsbacher & Hargreaves, 1988; Gernsbacher, Hargreaves & Beeman, 1989), and information structure. The latter will be highlighted in next section's overview on pronoun resolution research which is divided as follows: first, the interpretation of object pronouns and reflexives in children will be discussed. This is followed by adult L1 resolution of subject and object pronoun and lastly, I turn to L2 pronoun resolution.

Research on children's interpretation of object pronouns and reflexives shows that children acquire the correct interpretation of reflexive pronouns (e.g., *himself*) by three years of age, while their interpretation of object pronouns (e.g., *him*) is less adult-like (Grodzinsky & Kave, 1993; McKee, 1992). There are considerable cross-linguistic differences between children's performance. In some languages,

children correctly reject linking the object pronoun to the immediate subject referent most of the time, whereas in other languages children incorrectly link the object pronoun to the preceding subject referent (McKee, 1992). A variety of factors has been hypothesized to affect children's ability to correctly link the object pronoun to its referent. For instance, Ruigendijk, Friedmann, Novogrodsky & Balaban (2010) assume that referential function of the pronominal system plays an important role. The idea is that some pronominal systems are stricter in terms of their referential functions of pronouns. One prominent example is the comparison of the German and Dutch/English pronominal systems. While in Dutch and English, alongside a reflexive, an object pronoun in a prepositional phrase may be linked to the immediately preceding subject referent (as in (4a, 4b)), in German this function is solely executed by a reflexive (as in (4c)).

(4)   a. $Joey_1$ legde de bal naast $zichself_1$/hem_1.

       b. $Joey_1$ put the ball next to $himself_1$/him_1.

       c. $Joey_1$ legte den Ball neben $sich_1$/ihn_2.

New insight into what may be the underlying issue comes from research addressing the preceding discourse context. Spenader, Smits & Hendriks (2009) and van Rij, Hollebrandse & Hendriks (2016) have manipulated the discourse context preceding the object pronoun as to whether it contained a single referent or two possible referents. They found children to perform more adult-like in single referent contexts. In other words, children's performance improved when the previous context focused only one referent suggesting that children take the immediate context into account when processing the object pronoun. This was also reflected in their gaze data (van Rij et al., 2016). Focusing only one referent may have facilitated referent selection – as it shifted their attention to the focused referent – with children's limited working memory capacity and resulted in a more adult-like performance. The findings from these two studies indicate that manipulating information structure gives rise to the possibility that children process discourse information differently from adults which affects subsequent processing of pronouns. However, both of these studies have only looked into focusing referents

in a single referent context. The emerging question is then, if the same outcome applies to focusing a referent in contexts with more than one referent.

Turning now to adults' resolution of pronouns, in German, the personal subject pronoun *er* ("he") has been observed to be linked to the preceding subject/topical referent (e.g. Bader & Portele, 2019b; Bouma & Hopp, 2007; Colonna et al., 2012). One feature that distinguishes German from some other languages is that the demonstrative pronoun *der* (in more formal settings *dieser*) can be used anaphorically as well. These two subject pronouns have been found to prefer different referents. The difference for the two pronominal forms has been characterized in terms of opposite preferences for grammatical role as well as for information structure (Bosch, Katz & Umbach, 2007; Bosch, Rozario & Zhao, 2003; Comrie, 1997; Diessel, 1999; Kaiser, 2011b). The unstressed personal subject pronoun would be resolved towards the subject/topical referent, while the demonstrative is more likely to be linked to the object/non-topical referent. This idea has been captured by the form-specific approach (Kaiser & Trueswell, 2008) which states that pronoun resolution differs in the factors involved for the distinct referential forms. The form-specific approach is supported by studies targeting differences in referential preferences between pronouns. They show that referent preference for the two subject pronouns cannot be simply described as opposite preferences. More specifically, the personal pronoun is affected to a greater degree by grammatical role than the demonstrative pronoun, whereas the demonstrative pronoun is influenced more by information structure than the personal pronoun (e.g. Bader & Portele, 2019a; Kaiser, 2011c; Kaiser & Trueswell, 2008; Portele & Bader, 2016).

Unlike subject pronouns, resolution of object pronouns has not been as thoroughly investigated. Most research has been done on English, looking at the role of parallel structures. For instance, grammatical role parallelism describes that pronouns prefer referents of the same grammatical role. The first study in German to investigate grammatical role parallelism was that of Sauermann & Gagarina (2017). They found an effect of grammatical role parallelism, i.e. the subject referent was preferred with subject pronouns, while the object referent was preferred with object pronouns. However, they only considered gaze data which do not al-

ways reflect offline referent selection, i.e, final interpretation (cf. Blything, Järvikivi, Toth & Arnhold, 2021b; Schumacher, Dangl & Uzun, 2016; Schumacher, Roberts & Järvikivi, 2017). Moreover, they did not account for possible changes in information structure due to changes in word order (e.g. Frey, 2006; Sauermann, Hö hle, Chen & Järvikivi, 2011). Following Sauermann & Gagarina's experimental design, Abashidze, Gagarina & Bittner (2022) examined the influence of grammatical and positional parallelism during online and offline resolution of subject and object pronouns. Online preferences were guided by grammatical role (cf. Sauermann & Gagarina, 2017). Offline results showed a preference for the subject referent with the subject pronoun, while for the object pronoun, referent choice was at chance-level. Abashidze et al. explain their results in terms of topicality. They assume a topic bias alongside grammatical role parallelism. Since topics often coincide with the grammatical subject of a sentence (Reinhart, 1981), resolution of the subject pronoun would be straightforward as the topic bias as well as the grammatical role parallelism bias would select the same referent. In contrast, for the object pronoun, object and topical referent are separate referents. This would lead to the chance-level performance.

Most studies on pronoun resolution employ the unaccented form of a pronoun as the default intonation. According to Akmajian & Jackendoff (1970), accented or stressed pronouns receive a different interpretation than unaccented pronouns. As mentioned earlier, different theories link the use of certain referential forms to the referents' recoverablility (Ariel, 1988, 1990, 2001; Givón, 1983; Gundel et al., 1993). accented pronouns would indicate lower recoverablility than unstressed pronouns. More recent research suggests that resolution of unaccented and accented pronouns does not always lead to different referent preference. While Gillingham (2013) and Taylor, Stowe, Redeker & Hoeks (2013) assume that a reversal in preference depends on the plausibility of referents within the discourse context, Kehler, Kertz, Rohde & Elman (2008) and Mozuraitis & Heller (2017),following Rooth; Rooth's (1985; 1992) definition of focus, believe it depends on the alternative referents present in the sentence preceding the pronoun. Some structures would give access to alternatives more easily than others. In any case, it is

assumed that the deciding factor applies to both subject and object pronouns. For German, no research has investigated referent preference for accented pronouns thus far.

In L2 research on pronoun resolution, initially the target was to contrast pro-drop and non-pro-drop languages in L2 speakers. The main finding revealed highly proficient L2 speakers to diverge from L1 resolution patterns (e.g. Belletti, Bennati & Sorace, 2007; Okuma, 2011). When L1 and L2 were of the same type (i.e., both pro-drop, or both non-pro-drop) some research shows the gap in performance not to be reduced (Lozano, 2018; Polio, 1995), whereas others show there is a complex effect of L1 in L2 speakers, where online processing diverges but final referent selection to aligns with that of native speakers (Roberts, Gullberg & Indefrey, 2008). As for information structure effects, some research suggests L2 speakers rely on it to a greater extent than native speakers (e.g. Ellert et al., 2011; Patterson et al., 2017; Schimke & Colonna, 2016; Wilson, 2009), while other findings indicate L2 speakers to not use or have difficulties using information structural cues (e.g. Abashidze et al., 2022; Okuma, 2011). Nevertheless, there seems to be a common consensus that proficiency is a crucial factor for L2 speakers to become more native-like in their referential patterns (e.g. Ellert, 2010; Ellert et al., 2011; Lozano, 2018; Polio, 1995).

1.3   The Present Dissertation

The dissertation aims to expand on the existing research on pronoun resolution by explicitly addressing effects of information structure in discourse contexts. The aim is three-fold. The first (1) is to understand how information structure effects influence processing not only of pronouns but also the processing of the preceding referents. More specifically, I want to know whether information structure manipulations on referents triggers changes in the processing of pronouns and ultimately changes in referent selection. Conversely, do information structure manipulations on the pronoun affect referent selection? Secondly (2), I want to understand whether changes in information structure are processed differently in

9

the different groups under investigation. Particularly, I ask whether children and L2 speakers would benefit in their referent choices when changes in information structure shift their attention to one of the referents. Lastly (3), I want to find out the degree to which information structure affects the different types of pronouns.

There are many ways to manipulate information structure. Besides the use of syntactic structures such as it-clefts, or prosody, I make also use of German allowing changes in word order which can trigger changes in information structure as well (Frey, 2006). One factor that distinguishes the studies here from (most) other studies is the use of context, i.e., I do not look at the information structural manipulations in isolation, instead I embedded them in a context to render their use felicitous and more 'natural'. Additionally, I carefully consider and manipulate prosody in our experimental items. A factor that has been overlooked in pronoun research, especially with regards to information structure.

The present dissertation uses eye-tracking and behavioral data to test how different kinds of manipulations in information structure affect different types of pronouns, particularly, personal object pronouns and reflexives, as well as personal and demonstrative subject pronouns in three studies with different populations. The following three chapters are intended to be accessible as separate research articles and there is some overlap in the information they include. This overlap mostly is mostly to be found in the research overviews they include. Apart from these similarities, each chapter focuses on different pronouns, information structure manipulations, and researched population.

Chapter 2 investigates the effects of *it*-clefts on discourse referents as well as object pronouns and reflexives in German children and adults and how information structure processing can help explain why some children may struggle with their interpretation of object pronoun. Using eye-tracking, I examined the processing of not only the pronoun, but also the discourse preceding the pronoun, as well as the time window where participants made their interpretation choice. This allows us to get an insight of how the effect of information structure unfolds over time and what its role in the final interpretation is. This study aims to address specifically research questions (1) and (2).

Chapter 3 examines how a prosodic focus manipulation on the discourse referents affects their processing and how it subsequently influences processing of personal and demonstrative subject pronouns in German adults. Moreover, it tests whether 'prominence' of referents determines pronoun resolution. Similar to the study in Chapter 2, I used eye-tracking to access the time course of the information structure effect from the sentence preceding the pronoun until the pronoun sentence. This enables us to investigate effects of information structure, subjecthood, order of mention as well as the idea of 'prominence' (in the form of attention towards referents) during processing and for the referent selection. This study contributes to answering research questions (1) and (3).

Chapter 4 uses the same prosodic focus manipulation on the discourse referents as in Chapter 3, but examines referent selection for personal and demonstrative subject pronouns in second language (L2) learners. Further, it describes a prosodic information structure manipulation on pronouns, that is, the effect of unaccented and accented personal subject and object pronouns in German L1 and L2 speakers. Since in German, the demonstrative subject pronoun has been found to display opposite referent preferences to the personal subject pronoun, and accented pronouns can also show opposite preferences, using accented personal pronouns may give rise to a more detailed picture of the different levels of sensitivity the various pronoun types exhibit. Further, the use of the object pronoun allows us to contrast possible effects of role parallelism to position parallelism, since in German the object pronoun can occupy first- and second mention position. This study aims to provide answers to all research questions mentioned above.

Chapter 2

Focus Effect Unveils Children's Local Processing of Pronouns and Reflexives

Abstract

Studies on young children's comprehension have shown that children can experience problems interpreting object pronouns, even when reflexive interpretation is already adult-like. Compared to resolving reflexives, linking pronouns to a referent is considered a more "intensive" process, because it also involves non-syntactic factors like discourse context. This could explain why children experience more difficulties with pronouns than with reflexives. Using eye-tracking and a truth value judgement task, we investigated the effect of focus via *it*-clefts on the processing of reflexives and pronouns in German-speaking children and adults. I analyzed gaze data of three time segments: before, during and after the mention of the pronoun/reflexive. The cleft segment revealed similar processing of *it*-clefts in children and adults. In the subsequent reflexive/pronoun segment, clefts caused adults to pay overall more attention to the reflexive referent, while children fixated the clefted pronoun referent more. The third segment showed attention shifted to the incorrect referent when visual and linguistic stimuli did not align. The difference in focus effect, i.e., children attend the pronoun referent more, while adults pay more attention to the reflexive referent, helped uncover processing differences between children and adults. That is, unlike adults, children consider only the local discourse context during referential processing. I argue that these processing

differences cause children's interpretation difficulties. However, the offline data showed no effect of information structure, suggesting that whether the processing differences transfer to the final interpretation depends on the language-specific function of the pronoun system which may aid in restricting referential links.

2.1 Introduction

When it comes to pronoun resolution in children, a lot of the research has focused on interpretation of subject pronouns. Studies have investigated the role of subjecthood, order of mention, gender, and agentivity (e.g. Arnold et al., 2007; Blything, Iraola Azpiroz, Allen, Hert & Järvikivi, 2021a; Hickmann & Hendriks, 1999; Papadopoulou, Peristeri, Plemenou, Marinis & Tsimpli, 2015; Song & Fisher, 2005), semantic factors such as verb transitivity (e.g. Pyykkö nen, Matthews & Järvikivi, 2010), as well as syntactic and prosodic marking of information structure (e.g. Blything et al., 2021b; Järvikivi et al., 2014; Song & Fisher, 2007) during the resolution of subject pronouns.

As for object pronouns, the focus of child language research lies on what seems to be an asymmetry in the interpretation of object pronouns and reflexives. This research generally suggests that while children interpret reflexives accurately, they may still assign the incorrect referent to the object pronoun (Avrutin & Wexler, 1992; Grimshaw & Rosen, 1990; McKee, 1992; Philip & Coopmans, 1996; Sigurjónsdóttir & Hyams, 1992; Spenader et al., 2009).

In this paper, we aim to contribute to the clarification of what the underlying issue in children's pronoun processing and final interpretation is. Employing *it-*clefts to modulate the information structure of the preceding discourse context, we investigate the extent to which discourse information affects children's processing of reflexives and pronouns and whether it differs from adults'. Further, we include the analysis of the gaze data of three separate segments, to show that the use of the eye-tracking method should be extended to contribute to a better insight into the processing of the unfolding sentences.

Prior research shows that children acquire the correct interpretation of reflexive

pronouns by three years of age (Grodzinsky & Kave, 1993; McKee, 1992). Moreover, their performance in reflexive interpretation, such as in sentence 1a, is remarkably consistent regardless of the input language and regardless of the morphosyntactic status of reflexives, which varies across languages (Guasti, 2002). For instance, in languages like Spanish, Italian, and French, reflexives are expressed by a clitic form (Spanish/French: *se*, Italian: *si*), whereas in English a non-clitic form is used (*x-self*) (Guasti, 2002, 289). Dutch and German, in turn, have more than one reflexive form: a weak form (*zich/sich*) and a strong form (*zichzelf/sich selbst*) (Hendriks, 2014).

1. Monkey and bunny are in the bathtub.

    (a) Monkey is washing himself.

    (b) Monkey is washing him.

In contrast to reflexives, children's interpretation of object pronouns as in 1b is not always adult-like and can vary across languages. In a comparative study, McKee (1992) found a substantial difference between the performance of Italian-speaking children and English-speaking children. Italian-speaking children correctly rejected the anaphoric reading of object pronouns (i.e. linking the object pronoun to the immediate subject referent) 80-85% of the time, whereas English-speaking children correctly rejected it only 18-38% of the time. Similar performance as with Italian children has been found for Spanish (Baauw & Cuetos, 2003; Padilla, 1990), French (Hamann, 2002), Greek (Varlokosta, Karafoti & Karzi, 2000) and German (Ruigendijk et al., 2010). Languages that follow the English pattern (Grimshaw & Rosen, 1990; McKee, 1992) are Dutch (Philip & Coopmans, 1996; Spenader et al., 2009), Icelandic (Sigurjónsdóttir & Hyams, 1992), and Russian (Avrutin & Wexler, 1992).

Various theoretical accounts have been proposed to explain the asymmetry in the interpretation of reflexives and object pronouns, as well as why it is not observed for all languages.

Within the generative framework, a possible explanation for children's performance would be that while children have the required grammatical knowledge,

they lack the pragmatic knowledge (Chien & Wexler, 1990; Grodzinsky & Reinhart, 1993; Grodzinsky, Wexler, Chien, Marakovitz & Solomon, 1993; Thornton & Wexler, 1999), such as Rule I(Grodzinsky & Reinhart, 1993) to reject the incorrect interpretation of pronouns. For instance, in (5a) where the pronoun *him* as well as the reflexive *himself* can refer to *the boy*, Rule I would prefer the use of the reflexive over the use of the pronoun. The reason is that reflexives, depending on the linguistic framework, just require syntactic binding, co-indexation, or similar structural processes, whereas pronoun resolution is often affected by non-syntactic factors such as discourse context (van Rij et al., 2016; Spenader et al., 2009). Bergmann, Paulus & Fikkert (2012) suggest that the deviation between reflexive and pronoun resolution would imply that reflexive resolution would underlie a more "low-level, automatic" process, whereas pronoun resolution would be a more "intensive" process as it involves non-syntactic factors (Bergmann et al., 2012, p. 800). So, Rule I would prefer the low-cost processing of reflexives.

As for the cross-linguistic differences, Ruigendijk et al. (2010) assume that the asymmetry in acquisition between pronoun and reflexive interpretation is absent in some languages because these languages are more restricted in their object pronoun use. In example 5a, the reflexive *himself* and the pronoun *him* can refer to *the boy*. Unlike English however, in German 5b, only the reflexive *sich* can refer to *der Junge*. Therefore, for German-speaking children, pragmatic knowledge would not be needed to reject the incorrect referents for pronouns, as this would be already ruled out by the function of the pronominal system.

(5) a. The boy put the ball behind himself/him.

b. Der Junge legte den Ball hinter sich/ihn.

c. The boy enjoyed himself.

d. Der Junge amüsierte sich.

Thus, the pronominal system in German shows functionally more complementarity than English or Dutch. This idea is supported by Ruigendijk & Schumacher (2020)'s findings that in German, in contrast to English (Schumacher, 2005) and Dutch (Schumacher, Piñango, Ruigendijk & Avrutin, 2010), reflexives in coargu-

ment relations (see 5d, 5c) depend on the same interpretive mechanisms as reflexives in prepositional phrases (see 5b).

In contrast to the above accounts, usage-based approaches argue that children have yet to fully acquire the grammatical knowledge needed to be able to reject the coreferential interpretation. These accounts assume that the knowledge is gradually derived from the input (Matthews, Lieven, Theakston & Tomasello, 2009). Until this knowledge becomes robust, children's ability could diverge according to the frequency of individual lexical items (verbs and pronouns) and constructions being tested. This means the more familiar children would be with certain lexical items, the fewer mistakes they would make with reference resolution (Matthews et al., 2009). Matthews et al. (2009) tested this prediction in a truth-value judgment task (combined with a sentence production component). However, against this prediction, an inverse effect of verb frequency was found, where children's accuracy was higher with low frequency verbs than with high frequency ones, suggesting that the asymmetry might not be explainable from input frequency alone.

Another idea regarding the interpretation difficulties was put forward by van den Akker, Hoeks, Spenader & Hendriks (2012). In contrast to Matthews et al. (2009) who looked at the overall frequency of verbs, van den Akker et al. (2012) looked at the frequency of pronouns and reflexives occurring with certain types of verbs. More precisely, they explored whether children's (range 4;2–6;1) interpretation of reflexives and object pronouns would be sensitive to the semantic properties of the verbs by contrasting grooming verbs such as *wash* with regular transitive verbs such as *hit*. Grooming verbs tend to be more often "self-directed" than action verbs that tend to be more often "other-directed". Furthermore, Dutch has two reflexive forms, a simplex expression *zich* and a more complex reflexive form *zichzelf*, with self-directed events prefering *zich* and other-oriented events prefering *zichzelf* (van den Akker et al., 2012). The authors hypothesized that children who still have difficulties with the interpretation of pronouns and reflexives will rely on frequency information for their interpretation. Thus, they expected children to prefer the self-directed interpretation with grooming verbs and the other-directed interpretation with transitive verbs which would lead to more errors in

pronoun interpretation with grooming verbs and errors in reflexive interpretation with transitive verbs. They found for children performing poorly with both pronoun interpretation and with reflexive interpretation that mostly errors occurred for pronouns with grooming verbs, while reflexive interpretation errors occurred with transitive verbs. This suggests that the combined frequency effect of verb type and pronoun type (reflexive/object pronoun) indeed played a role during pronoun interpretation for younger children. However, for children that performed well with reflexives but made errors with pronouns, no difference between the two verb types was found. Moreover, this account cannot explain the cross-linguistic differences found for pronoun interpretation.

Clackson, Felser & Clahsen (2011) investigated older English-speaking children (6-9 years) and adults' processing of reflexives in comparison to object pronouns using contexts in which the two possible referents matched in gender or did not. For pronouns, children and adults looked more at the incorrect referent when both referents were matched in gender than when they were not. Similarly, for reflexives, children and adults fixated the incorrect referent more when it matched in gender. However, this effect was more pronounced in children than in adults. For the offline interpretation, adults' performance was at ceiling level in all conditions, whereas for children accuracy was high for reflexives, but for pronouns, accuracy decreased when referents were matched in gender. Finally, Bergmann et al. (2012) tested online and offline pronoun resolution in an eye tracking experiment and picture selection task (same items). They found that Dutch-speaking children at the age of 4 fixated the correct referent, but still failed to choose the correct referent in the offline task. Bergmann et al. (2012) suggest this discrepancy could have emerged from task-inherent demands overexerting children's cognitive resources (cf. Conroy, Takahashi, Lidz & Phillips, 2009). This explanation suggests that children's performance during interpretation solely depends on the cognitive demands of the task.

While most research has focused on children's difficulty in interpreting object pronouns, some researchers have approached the issue from a different angle. For instance, Hendriks, Banga, van Rij, Cannizzaro & Hoeks (2011) looked into adult's

processing of object pronouns and reflexives in Dutch in order to better comprehend children's difficulties. They measured adults' accuracy, response times and eye movements during a picture verification task. As expected, accuracy was high with object pronouns and reflexives. The more interesting results came from the response times and eye movements measures. The authors found adults to have longer response times with object pronouns than with reflexives when there were two possible referents compared to when there was only a single referent in the preceding context. They interpreted the longer response times to reflect higher processing efforts for pronouns. For the eye movements, they discovered shorter fixation times towards the correct referent for pronouns than for reflexives, as well as participants taking longer to fixate the correct referent for pronoun than for reflexives regardless of the previous context. While taking longer to identify the correct referent for the pronoun, the shorter fixation times also seem to indicate that participants found the correct referent for the pronoun less probable. Overall, these results seem to point to remnants of the interpretation difficulties found for object pronouns in children's language processing to still be present to a small extent in adults.

Another study taking adults' object pronoun processing into account is the one by Vogelzang, Hendriks & van Rijn (2016). They too compared processing of object pronouns and reflexives in Dutch adults using the measure of pupil dilation. Pupil dilation has been found to be an effective tool measuring processing effort during sentence processing (e.g. Borghini & Hazan, 2018; Just & Carpenter, 1993). Similar to Hendriks et al. (2011) their results indicate possible remnants of pronoun interpretation difficulties as pupil dilation was larger for object pronouns than for reflexives which would reflect more processing effort for pronouns than reflexives. Lastly, Vogelzang, Hert & Ruigendijk (2021b) explored object and reflexive processing in German adults by means of pupil dilation in order to compare processing of adult German with Dutch speakers. Their gaze data showed no differences in processing efforts between object pronouns and reflexives in German. Object pronouns in Dutch may require greater processing effort, possibly because they are functionally more ambiguous compare to German pronouns. These findings

therefore might suggest that children's interpretation difficulties have the same underlying source, namely increased processing efforts for object pronouns.

Taken together, these studies show that adults' data on pronoun processing and interpretation should be included not only to compare it to children's data in order to investigate when children become adult-like, but also to gain more insight into where these interpretation difficulties stem from and how they differ cross-linguistically.

Regarding the interpretation difficulties, it may be useful to take discourse context and the explicit use of information structure into account. While some studies found children's processing of reflexives and pronouns to be similar to that of adults (e.g. Bergmann et al., 2012; Clackson et al., 2011; Sekerina, Stromswold & Hestvik, 2004), there is also evidence that children's processing is more context-dependent than adults' and therefore processing can vary with respect to how similar children and adults are (van Rij et al., 2016). Information structure cues for focus marking, such as clefts, for example *It is the penguin that hits the monkey*, have been argued to increase antecedent prominence by placing focus, or contrast, on the clefted item. Further, clefting can direct the reader's or listener's attention to the clefted entity (Klin, Weingartner, Guzmán & Levine, 2004). In addition, when focus is marked by means of clefting, it can often indicate an exhaustive interpretation (Büring & Križ, 2013; Drenhaus, Zimmermann & Vasishth, 2011; Kiss, 1998; Krifka, 2008).

In a number of experiments with adult participants, Birch, Albrecht & Myers (2000) investigated the influence of syntactic focusing structures on discourse processing during reading. Their results suggest that clefting strengthens the memory representation of clefted constituents compared to non-clefted constituents. Similarly, Norberg & Fraundorf (2021) explored the effects of different contrastive focus markers during reading and found clefts, but not other markers, to aid in retaining focused words (Kember, Choi, Yu & Cutler, 2021, also see). Foraker & McElree (2007) conducted three experiments exploring the effects of syntactic clefting in adults. In experiment 1 and 3, cleft conditions showed higher accuracy compared to the non-cleft condition indicating that clefting increases the availability of

antecedent representation. Experiment 2 measured eye fixations during reading. The results showed an effect in later eye-tracking measures: compared to clefted conditions, non-clefted conditions showed longer regression path times, second-pass times, as well as total reading times. The authors therefore assume that clefting facilitates antecedent integration into discourse. Altogether, their results indicate that clefting affects antecedent representation by making it more distinctive in working memory.

Similarly, offline tasks that contrasted contexts whether one or two referents were included showed that children's performance on pronoun interpretation increased significantly in the single-referent context when compared to the two-referent context (van Rij et al., 2016; Spenader et al., 2009). This was also reflected in the gaze data (van Rij et al., 2016). While the two-referent context focused two referents, the single-referent context focused only one referent, thus there was less competition for the pronoun's referent. Focusing only one referent may have aided with children's limited working capacity and led to a more adult-like performance for object pronoun interpretation.

The effect of syntactic focus marking via clefting has not been investigated yet for object pronouns in German. For Spanish object clitics, clefting has been found not to significantly change referent choice when compared to the baseline condition (De la Fuente & Hemforth, 2013). However, the authors used pronouns in ambiguous contexts and only looked at their final interpretation. Moreover, the cleft constructions were presented without a preceding context which does not correspond to the function of *it*-clefts, namely to establish contrastive focus or exhaustive interpretation with something from the previous discourse (Destruel & Velleman, 2014; Kiss, 1998). Consequently, if the cleft is introduced without any preceding context this function could not be fulfilled.

For subject pronouns, Järvikivi et al. (2014) found German-speaking children to be sensitive to the cleft structures in their study. Unlike for adults, for children clefting the subject referent increased attention towards the subject referent. Yet clefting the object referent did not have the same effect. Thus, for children, clefting seemed to only draw more attention to the preferred subject referent in

the interpretation of subject pronouns. However, Blything et al. (2021b) found for English-speaking children and adults that syntactic focus marking via clefting did not have an effect on looks towards a referent when prosodic focus marking was already present. Offline interpretation revealed a consistent subject (first-mention) preference for the subject pronoun regardless of the previous online preference, in line with the above memory-based accounts.

If clefting is seen as a "memory aid", then these results could also suggest that clefting highlights the referent holding the same grammatical role as the subject pronoun, making it easier to retrieve from memory. This may also explain the differences found in children and adults in Järvikivi et al. (2014)'s study. Children may need to rely more on a "memory aid" than adults because of their more limited working memory capacity (Cowan, 2010; Kharitonova, Winter & Sheridan, 2015). Thus, clefting may help shifting attention to the clefted referent which may make the referent more distinctive in working memory. This in turn could benefit children's pronoun processing since a clefted referent may be more easily retrieved from memory than one that is not clefted.

Also, beyond the use of clefts, the context that includes the possible referents seems to play a major role in children's behaviour and performance during processing and interpretation of pronouns. Children show sensitivity to changes in the discourse context as they can affect their processing. Moreover, when these changes shift attention to the intended referent, they can indeed improve children's performance during pronoun interpretation (Clackson et al., 2011; van Rij et al., 2016; Spenader et al., 2009). Additionally, it seems that the preceding context cannot only modulated the processing of pronouns, but may also affect processing of reflexives (Clackson et al., 2011). Yet, studies have not analyzed processing of the context itself. However, doing so may be useful for understanding possible differences in processing that may emerge from changes in the discourse context. Moreover, it would also shed light on whether children and adults process preceding contexts differently which could transfer to processing the pronominal elements and ultimately may also affect their interpretation.

By addressing the question whether information structure in the form of clefts

can affect how reflexives and object pronouns are processed when the object pronoun's referent is focused, we aim to investigate the research gaps and issues mentioned above. More specifically, these are: (1) Whether there are differences in children's and adult's processing of discourse context with respect to changes in the information structure; (2) whether changes in information structure affect subsequent processing of pronouns and reflexives and whether children's processing of pronouns benefits from discourse context focusing the intended referent; (3) and whether any of the potential differences between adults and children and/or reflexive and pronoun processing can help explain the cross-linguistic differences found in children.

2.2 Current study

In this visual world eye tracking experiment, we used *it*-clefts to focus the object pronoun's referent. Children and adults listened to mini stories containing four animal characters (see 6 for an example) while we tracked their gaze to pictures displaying these four animals. We manipulated whether (i) participants heard a no cleft (see 6a) or a cleft (see 6b) sentence; (ii) whether the sentence contained a reflexive or pronoun (see bold words in 6c); and (iii) whether the picture displayed after the story was played was congruent (e.g. picture displayed self-directed action, the story contained a reflexive) or incongruent (e.g. picture displayed a self-directed action, story contained a pronoun). We also collected their offline decision regarding reflexive and pronoun interpretation. We asked how clefting would affect processing of the possible referents as well as the pronominal elements.

(6)   Der Hund, der Igel, der Löwe und der Affe sind gemeinsam auf einer Geburtstagsfeier.
'The dog, the hedgehog, the lion and the monkey are together at a birthday party.'
(a) Der Affe bewirft den Igel mit Konfetti, (*no cleft condition*)
'The monkey throws confetti at the hedgehog'
(b) Es ist der Affe, der den Igel mit Konfetti bewirft, (*cleft condition*)

'It is the monkey who throws confetti at the hedgehog'

(c) während der Igel **sich/ihn** vor den Luftballons fotografiert.

'while the hedgehog photographs **himself/him** in front of the balloons.'

In order to investigate the effects of information structure on reflexive and pronoun processing in children and adults,we chose German, because it has been identified in previous research as a language where children do not display difficulties interpreting object pronouns (Ruigendijk et al., 2010). This allowed us to rule out possible influences coming from expected differences in children's mastery of reflexives and pronouns. Therefore, if we do find differences between cleft and no cleft conditions in processing and/or final interpretation in our results, we can assume that these are due to the focus manipulation and that this effect appears regardless of whether or not children have mastered pronoun interpretation. Finding differences in processing and/or final interpretation between no cleft and cleft conditions would suggest that shifting attention to the intended referent may be beneficial for reference processing and/or interpretation. If children differ from adults, it would suggest that the degree of discourse context information used varies between these two groups during reference processing. If German-speaking children rely more on discourse information than adults, even though they generally do not struggle with pronoun interpretation, it may be beneficial to apply this to languages where children experience pronoun interpretation difficulties to aid in their interpretation process. If the effect of information structure unfolds similarly for reflexives and pronouns (i.e. attention is directed towards the same referent), then this might indicate that information from the discourse context is preferred over syntactic/pragmatic (e.g. Rule I) information of the reflexive/pronoun.

Note that in the current study, the object pronoun and the clefted referent do not hold the same grammatical role. However, we do not expect grammatical role to influence referent preference, since we do not use ambiguous object pronouns (Sauermann & Gagarina, 2017, cf.).

Moreover, in order to gain better insight into the processing throughout the unfolding sentences, the analysis of the gaze data includes three separate segments:

(i) **Segment 1** is the segment preceding the referential expression which included the syntactic manipulation (e.g. 6a/b above); (ii) **Segment 2** which included the reflexive/pronoun (e.g. 6c); (iii) **Segment 3** which started with the appearance of the picture displayed after the end of the story and continued until participants' response. This allowed us to ensure that clefting really results in more looks – increased attention – to the clefted referent compared to no cleft sentences. It further allowed us to explore whether the effect carries over to the pronoun/reflexive segment and how that affects the processing of pronouns and reflexives. And finally, we could also analyze the gaze behavior during the decision process.

If clefting affects reference processing, we would expect children to look more to a clefted referent compared to the condition where the referent is not clefted. I expect an effect of clefting, i.e. both adults and children to look more towards the pronoun's referent when that referent is clefted than when it is not clefted during Segment 1. Upon hearing the pronoun, we expect children as well as adults to look more towards the pronoun's referent in the cleft condition compared to the simple no cleft condition during Segment 2. If this were indeed be the case, it would mean that focusing the pronoun's intended referent via clefts would draw more attention to the referent. On the other hand, if the pronoun's intended referent were not to receive more fixations, then focusing via clefts would not lead to more attention.

If clefting leads to a shift in attention to referents, then upon hearing the reflexive, children as well as adults may look more towards the pronoun referent than to the reflexive referent in the cleft condition, since it was always the pronoun's referent that was clefted. This was also the case in Clackson et al. (2011) where changes in the prior context affected both object pronouns and reflexives, as well as in Hert, Järvikivi & Arnhold (n.d., Chapter 3 in this thesis), where the object referent received more attention when it was focused in the preceding discourse. However, other studies reported that clefts did only have an effect on the preferred referent (Blything et al., 2021b; Järvikivi et al., 2014), which, in our study in the case of reflexives, is the immediately preceding subject referent. Therefore, we might not see a difference for reflexive processing in the cleft and no cleft conditions.

For the offline interpretation, Colonna, Schimke & Hemforth (2015) found clefted

24

referents to be chosen more in German for intersentential resolution of subject pronouns. Similarly, Hert, Järvikivi, and Arnhold (submitted) found subject referent choices in German to decrease when the object referent was focused in the prior context (see also Cowles, Walenski & Kluender, 2007; Foraker & McElree, 2007; Norberg & Fraundorf, 2021). On the other hand, Blything et al. (2021b) showed no difference in referential choice for clefted referents in English (see also Colonna et al., 2012; De la Fuente & Hemforth, 2013; Kaiser, 2011a; Patterson et al., 2017). Unlike in the previous studies, in our study pronouns were not ambiguous, since the context (story in combination with picture) for the reflexives and pronouns only allows one correct referent. Therefore, we do not predict a difference between the cleft and the no cleft conditions with respect to referent choice, neither for adults nor for children.

2.3 Methods

2.3.1 Participants

Altogether, we collected data from 18 monolingual German children (range 4;1-6;6, mean age 4;10) from Konstanz and Lauenbrück kindergartens and from 51 native German-speaking adults (range 20-33, mean age 24) from the universities of Konstanz and Oldenburg. Adult participants gave written consent while parental consent for children was obtained prior to the testing session. Children received a certificate and a small gift and adults received monetary compensation for their participation. This study was approved by the Research Ethics Board 2 of the University of Alberta (study ID Pro00090878).

2.3.2 Materials

Thirty-two mini-stories (see example 6) were created for a 2x2x2 design. We manipulated (i) the referring expression in the test sentence, a pronoun or a reflexive; (ii) whether the test sentence contained a cleft or not; and (iii) for the offline data only, the picture presented on the screen, showing an other-oriented action (congruent with a pronoun sentence, but incongruent with a reflexive sentence) or a

self-oriented action (incongruent with a pronoun sentence, but congruent a reflexive sentence). We selected 16 verbs; each verb was used twice, for one self-oriented and one other-oriented picture. We added two practice items to familiarize participants with the task. Each story introduced four animal characters in the first sentence (see 7a). The following sentence contained either a cleft (see 7c) or not (see 7b) in the main clause and included only two of the four previously introduced animals. The clefted entity was always the subject of the main clause and the pro- noun's referent in the subordinate clause. The subordinate clause contained either a reflexive or a pronoun. All animal characters were masculine in gender to avoid number ambiguity from the feminine object pronoun *sie* 'she/they'.

(7)   a. Der Hund, der Igel, der Löwe und der Affe sind gemeinsam auf einer Geburtstagsfeier.
'The dog, the hedgehog, the lion, and the monkey are together at a birthday party.'

b. Der Affe bewirft den Igel mit Konfetti, während der Igel **sich/ihn** vor den Luftballons fotografiert.
'The monkey throws confetti at the hedgehog while the hedgehog photographs **himself/him** in front of the balloons.'

c. Es ist der Affe, der den Igel mit Konfetti bewirft, während der Igel **sich/ihn** vor den Luftballons fotografiert.
'It is the monkey who throws confetti at the hedgehog while the hedgehog photographs **himself/him** in front of the balloons.'

The stories were recorded in a sound treated chamber by a prosody researcher and native speaker of German. The clefted referent also received prosodic prominence. This was a natural way to focus the clefted antecedent in accordance with the typical prosodic realization of focus in German (Féry, 2017) (compare Figure 2.1a, where *Affe* 'monkey' is realized with a rising accent typical of prenuclear material and Figure 2.1b where it carries a high/falling accent typically found on focused constituents).

A picture showing all four animals in 'neutral' poses (see Figure 2.2a) accom-

Figure 2.1: *Example for prosodic contours and waveforms for no cleft and cleft sentences.*

(a) No cleft sentence

(b) Cleft sentence
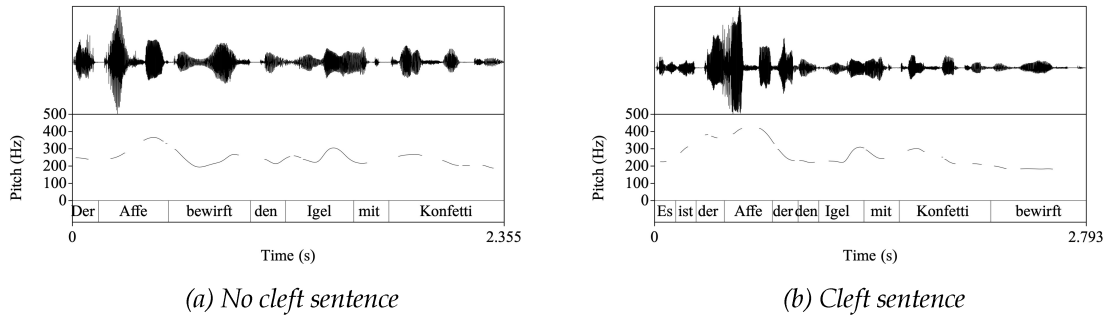
panied each mini-story. After the mini-story finished playing the picture was changed to one depicting the two animals that were mentioned in the last sentence 'in action', while the other two animals remained in 'neutral poses' (see Figure 2.2b). The last sentence (see 7b and 7c for example, as well as 6c above) was either congruent or incongruent with the picture. We chose to change the picture after the story was played, so that participants would not be biased by the visual scene when making their decision. In addition to gaze data, we also measured accuracy data. This measure started from the changed picture and carried on until participants responded with a button press.

Figure 2.2: *Example of visual stimulus displayed during eye tracking.*

(a) Picture shown while story was playing.

(b) Picture shown after story finished playing.

27
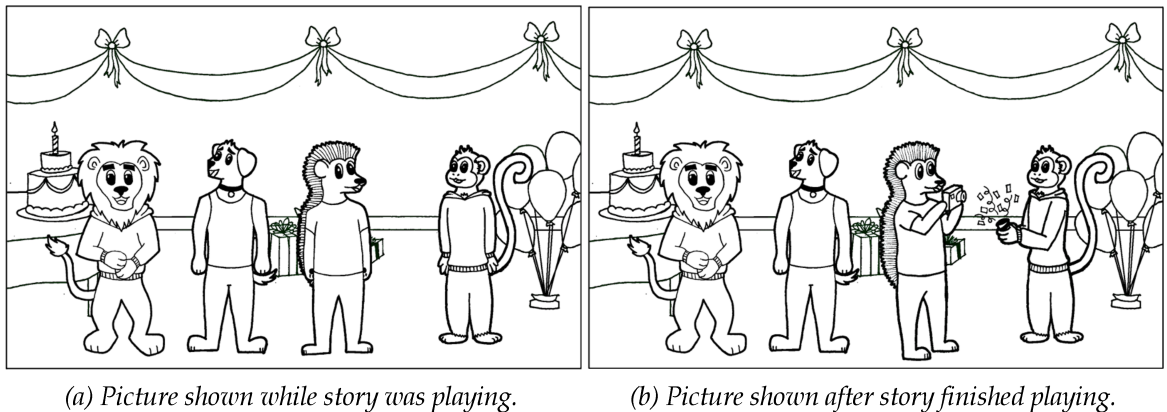
2.3.3 Procedure

The participants were instructed to listen to the stories and look at the pictures on the computer screen. They were told that the picture on the screen would change after each story and that in the truth value judgement task they had to decide whether the picture matched the story or not. The eye tracking experiment started with a 9-point calibration which was followed by two example trials. The order of the experimental items was randomized. An SR Research Eyelink Portable Duo eye tracker was used to record participants' eye movements. Registration was monocular (right eye) with a 500 Hz sampling rate.

2.4 Results

2.4.1 Truth Value Judgement Task

I ran generalized linear mixed-effects models – using R studio (version 4.0.2, R Core Team, 2020) *lme4* package (version 1.1-27.1, Bates, Mächler, Bolker & Walker, 2015) – to analyze the accuracy data. The binomial dependent variable was *Accuracy* on picture matching (accuracy was 1, if the picture was correctly matched to the sentence, e.g. a picture displaying a self-directed action was matched to a sentence containing a reflexive), independent predictors were *referential expression* (reflexive, pronoun), *congruency* (congruent picture, incongruent picture), and *Group* (adult, child), as well as a four-way interaction among these factors. We also included random slopes in the models– a by-subject random slope for congruency and a by-item random slope for congruency. One by one, factors were excluded to assess whether they contributed to the model's fit. A significant contribution was assessed by comparing the Akaike information criterion (AIC) score (Akaike, 1974) of the new model, which excluded the fixed factor under examination, with that of the previous model, which included the fixed factor but was otherwise identical. A decrease of at least 2 in the AIC scores indicates that the inclusion of a factor significantly improves the goodness of fit of the model. Similarly, the necessity of including random slopes was assessed. Both random slopes remained in the model as they improved the model's fit. The best fit model (table 2.1) did

not contain *condition* (cleft, no cleft) nor *referential expression* (reflexive, pronoun) as predictors since they did not significantly improve the model's fit.

Fixed effects:	Estimate	Std. Error	z value	p-value
(Intercept)	3.5532	0.3510	10.124	< 0.0001
congruencyincongruent	-1.3908	0.5393	-2.579	0.00991
GroupChild	-1.9029	0.4891	-3.890	0.00010
congruencyincongruent:GroupChild	1.9658	0.8132	-2.418	0.015629

Table 2.1: Model summary for Truth Value Judgement task.

Figure 2.3 visualizes participants' performance. It shows that in the incongruent condition accuracy was lower than in the congruent condition, as well as adults performing better than children.
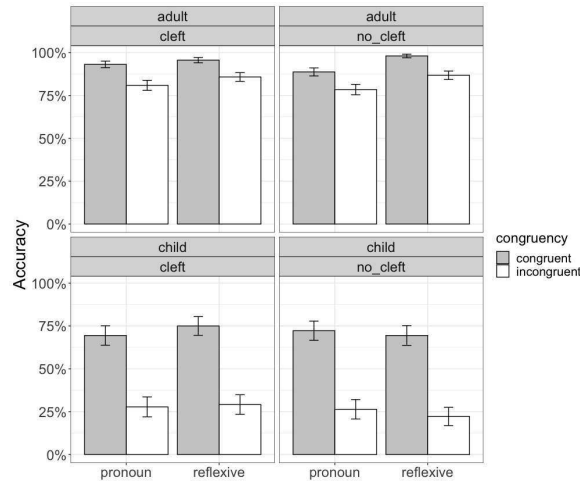
Figure 2.3: Adults' and children's performance in truth value judgement task comparing pronouns and reflexives in congruency and syntactic conditions, plotted with standard error bars.

Table 2.1 shows the model's predictions for the offline data. From the estimates we can infer that accuracy was lower for incongruent than for congruent conditions. Children's performance was lower than that of adults. The interaction between *congruency* and *Group* was significant. A comparison of the multiple levels of the interaction term revealed that, as expected, children performed similarly with pronouns and reflexives. This suggests that for children there is no difference between the interpretation of reflexives and pronouns, which is in line with previous findings for German-speaking children (Ruigendijk et al., 2010).

However, performance – specifically in the incongruent condition – was relatively low for both pronouns and reflexives, suggesting unadult-like performance (van Rij, van Rijn & Hendriks, 2010, cf.). Surprisingly, congruency was not only significant for children but for adults as well. Both groups performed worse in the incongruent condition, with children's performance being relatively poorer than adults'. For children, this could indicate "yes"-bias (Chien & Wexler, 1990, cf.,), but perhaps for adults as well.

2.4.2 Gaze Data

The gaze data were pre-processed using the VWPre package (Porretta, Kyröläinen, Rij, Järvikivi, van Rij & Järvikivi, 2018). We used 20 ms bins, downsampling the data to 50 HZ (each bin contained 10 samples). For the statistical analysis of the gaze data, we ran Generalised Additive Mixed Models (GAMM) using the package mgcv (version 1.8-38 Wood, 2017). GAMMs are suitable for handling time series data such as eye tracking data, because they do not entail averaging over time windows or assume that relationships between continuous variables are linear (van Rij et al., 2016). GAMMs allow for the inclusion of non-linear relationships using smooths (van Rij, Vaci, Wurm & Feldman, 2020a; Wieling, 2018; Wood, 2017). Analyzing fixation proportions using GAMMs is not recommended (Porretta et al., 2018; van Rij et al., 2016), so we transformed the data to e-logits using the *transform to elogit* function of the VWPre package. This function symmetrically distributes values around zero and yields an unbounded measure for analysis (Barr, 2008). For Segment 1, the response variable for the GAMM analysis was subject preference looks, calculated by subtracting the e-logit looks to the object character from the e-logit looks to the subject character. For Segment 2 and Segment 3, the response variable was reflexive referent preference looks, calculated by subtracting the e-logit looks to the pronoun referent character from the e-logit looks to the reflexive referent character.

With GAMMs, smooth functions are used to fit linear and nonlinear patterns of the data. A smooth function is made up of a weighted sum of a set of base functions that vary in shape (van Rij et al., 2016). In addition to account for variation in

participants and items via random slopes and intercepts like in linear mixed effects modelling, GAMMs allow for random smooths. To our models, we added by-participant and by-item random intercepts and slopes to time, as well as a by-event (each unique participant-item response) intercepts random smooth to the effect of time.

To determine the effects of the fixed factors, we followed a backward-fitting model comparison. In addition to the model summary output, we used the model comparison results – the *compareML* function of the itsadug package compares the Maximum Likelihood (ML) score of model variants – and visualization of the model to choose the best fit model (see e.g. van Rij et al., 2020a).

I analyzed three time windows: (i) the syntactic manipulation (cleft or no cleft) containing the two possible referents, **Segment 1**, (ii) the segment containing the pronominal element (object pronoun or reflexive), **Segment 2**, and (iii) the segment where the displayed picture changed and participants had to decide whether the new picture was congruent or incongruent with the previous sentence, **Segment 3**. Segment 1 was set to the onset of the cleft/subject continuing until the end of the of the sentence (approximately 2500 ms). Segment 2 was set to 200 ms prior to the onset of the reflexive/object pronoun (0 ms) continuing until the end of the pronoun sentence (approximately 3820 ms). Segment 3 was set to the beginning of the picture change until participants' button press.

Segment 1

The gaze data for Segment 1 showed the proportion of looks towards the subject and object referent in the cleft and no cleft condition for adults and for children (Figure 2.4).

For adults, both conditions showed an initial increase in looks towards the subject. For the cleft condition the proportion of looks was slightly higher compared to the no cleft condition. The start and the peak of looks towards the subject were relatively delayed in the cleft condition which was due to the subject being mention after the *It is* part of the cleft when compared to the no cleft condition which immediately started with the subject. Looks to the object increased in the second

Figure 2.4: Average proportion of looks in Segment 1 for adults and children in cleft and no cleft conditions. First dashed line marks the average onset of the clefted subject referent (cleft condition only), other dashed line marks the average onset of the object referent.
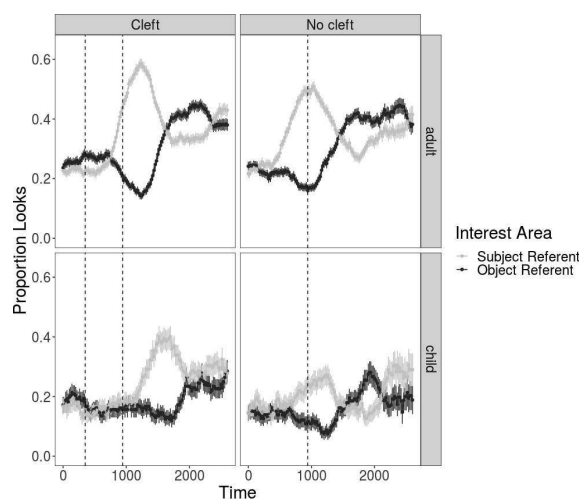
half of the segment, and while they decreased towards the end in the cleft condition, they remained constant longer in the no cleft condition before they decreased at the very end as well. The differences between the conditions could mean that clefts not only enhance attention towards the subject referent, but that their use leads adults to predict that the clefted referent will be mentioned again as looks to the subject referent towards the end of the segment increase earlier in the cleft than in the no cleft condition.

For children, looks to the subject increase in both conditions resulting in more looks to the subject than the object referent. The peak is higher in the cleft than in the no cleft condition. In the no cleft condition, the looks to the subject decrease approximately halfway through the segment, resulting in more looks to the object than the subject, but rise again towards the end. In the cleft condition, even after the peak the proportion of looks to the subject remains higher than the proportion of looks to the object. The difference between conditions could mean that clefts enhance attention towards the subject referent. In contrast to adults, children look more to the subject referent at the end of the segment. This could stem from a delay in children's (sentence) processing. As can be seen that in both conditions, the peak in looks for the subject referent emerges later than for adults.

The model for the first segment included subject advantage looks (elogits of

looks to the subject - looks to the object) as the dependent variable, as well as a combined factor of *Condition* (cleft, no cleft) and *Group* (adult, child) which was included as a **parametric term**, as well as smooth terms for the interaction between the combined predictor with *Time*. Further, the models included random intercepts for *Event* (a unique combination of *Subject* and *Trial*), and random smooths for *Subject* and *Item* to account for individual variation between participants. We also accounted for autocorrelation by adding an AR1 (autoregressive) parameter to the model which specified the autocorrelation coefficient (*rho* = 0.9) and the starting point for each time series (Baayen, van Rij, de Cat & Wood, 2018). Table 2.2 shows the model summary of the inferential statistics for the optimum-fit model for Segment 1. For the parametric terms (time course is not taken into account), the *p*-value reveals whether a *Condition* and *Group* was significantly different from the reference level *Cleft adult* which is represented as the intercept. Positive *Estimate* values indicate more looks to the subject referent overall, while negative values signal more looks to the object referent. The model predicted only the intercept to be significant, other combined factor of *Condition* and *Group* were not significantly different from the intercept when averaged over the whole segment.

For the smooth terms, the *p*-value indicates whether or not a smooth is significantly different from 0. The *edf* value represents the number of effective degrees of freedom which estimates the number of parameters needed for the smooth, and reflects the degree to which the pattern is (non)linear. A value of 1 corresponds to a linear pattern while a value greater than 1 suggests a more complex non-linear pattern. For Segment 1, all smooth terms, except for the two *No Cleft* conditions, were significantly different from 0 and non-linear (edf > 1).

To visualize the GAMM's predictions and understand how the effects contribute to the participants' attention towards the possible referents as well as how they unfold over time, differences among these conditions were plotted with the itsadug package (version 2.3, van Rij, Wieling, Baayen & Van Rijn, 2020b). These plots are needed for determining whether smooths of each experimental condition significantly differed from each other. On the y-axis are the estimated differences between the looks to the subject referent and looks to the object referent. A positive

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.4274	0.1828	2.3375	0.0194
ConditionGroupCleft_child	-0.0933	0.3125	-0.2985	0.7653
ConditionGroupNo cleft_adult	-0.1263	0.0805	-1.5682	0.1168
ConditionGroupNo cleft_child	-0.2787	0.3159	-0.8822	0.3777
B. smooth terms	edf	Ref.df	F-value	p-value
s(Time):ConditionGroupCleft_adult	7.8403	8.0125	5.6156	< 0.0001
s(Time):ConditionGroupCleft_child	8.3018	8.8529	41.8698	< 0.0001
s(Time):ConditionGroupNo cleft_adult	3.1951	3.4641	0.9316	0.3963
s(Time):ConditionGroupNo cleft_child	1.0132	1.0152	0.2309	0.6354
s(Event)	1646.9237	1980.0000	15.7517	< 0.0001
s(Time,Subject)	479.2067	696.0000	11.0096	< 0.0001
s(Time,Item)	276.3852	318.0000	388.0077	< 0.0001

Table 2.2: Model summary for gaze data of Segment 1

value indicates more looks towards the subject referent than towards the object referent for the condition first mentioned in the title, while a negative value indicates more looks towards the object. On the x-axis, 0 ms is the onset of the no cleft/cleft. Significant differences in looks to the two target referents are marked with dashed lines. Only plots that contain significant time windows are presented here, for all other plots check `https://osf.io/g5vhq/`.

Figure 2.5 shows the significant differences between different combinations of Group (adult, children) and Condition (cleft, no cleft). For Figure 2.5a, the first significant time window (-100-31 ms) will be disregarded since it started before the sentence onset. The next significant time window (215-976 ms) reveals significantly more looks to the object referent than the subject referent in cleft condition compared to no cleft condition. The following significant time window (1055-1712 ms) shows significantly more looks to the subject referent in the cleft condition than in the no cleft condition. This is followed by another shorter time window (1869-2158 ms) which shows more looks to the object referent in clefts than in no cleft condition. The final significant time window (2342 ms - end of segment) captures more looks to the subject referent for clefts than in no cleft condition. Figure 2.5b depicts the same comparison between cleft and no cleft sentences, but for children. The first significant time window (792-1239 ms) shows more looks to the object referent than the subject referent in cleft condition compared to no cleft condition. The fol-
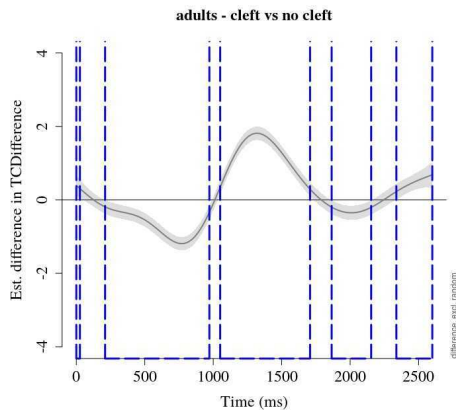
(a) Difference in looks between cleft and no cleft for adults
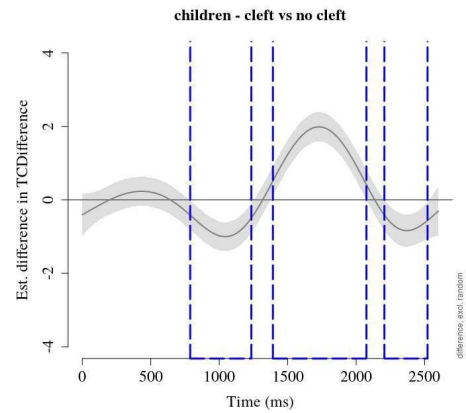
(b) Difference in looks between cleft and no cleft for children
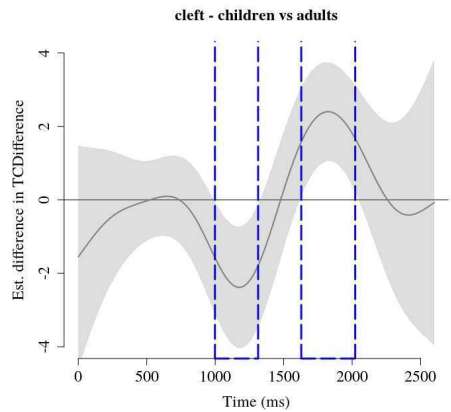
(c) Difference in looks between children and adults for cleft

Figure 2.5: Difference plots for Segment 1 comparing Group (children, adults) and Condition (cleft, no cleft). Shaded areas represent 95% confidence interval. Horizontal black line indicates zero effect. Dashed lines mark significant time windows.

lowing time window (1296-2053 ms) reveals significantly more looks to the subject referent in clefts than in no cleft condition. And the final significant time window (2211 ms-end of segment) depicts more looks to the object referent for the cleft condition compared to the no cleft condition. Figure 2.5c compares gaze patterns of children and adults in the cleft condition. The first significant time window (1003-1318 ms) shows more looks to the object referent for children compared to adults. The following significant time window (1633-2027 ms) depicts more looks to the subject referent for children than for adults. For the comparison of gaze patterns between children and adults for the no cleft condition, there were no significant

(a) pronouns   *(b) reflexives*
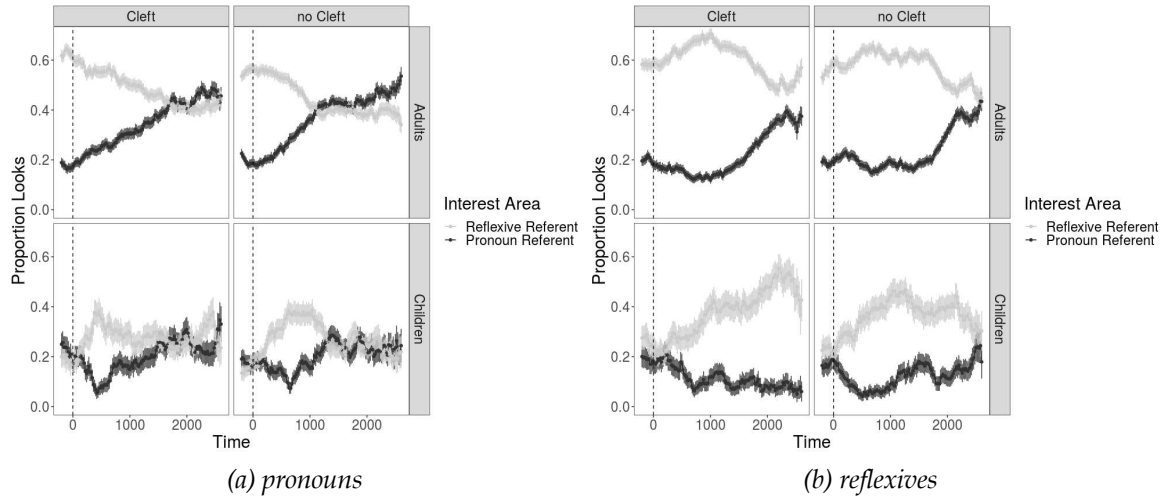
Figure 2.6: Average Proportion of looks for Segment 2. Adults and children in cleft and no cleft condition. Dashed line indicates onset of reflexive/pronoun.

time windows.

In sum, the results of the gaze data in Segment 1 showed that the overall gaze pattern for the comparison of the cleft and no cleft conditions was similar in adults (Figure 2.5a) and children (Figure 2.5b): first more looks to the object referent, followed by more looks to the subject referent in the cleft condition than for the no cleft condition. The direct comparison between children and adults in the cleft condition revealed that children's peaks in looks in the first (more looks to object referent) and second (more looks too subject referent) time window were later than for adults.

Segment 2

The raw gaze data for Segment 2 for pronouns (Figure 2.6a) and for reflexives (Figure 2.6b) shows overall attention towards the two referents across cleft and no cleft condition for adults and children. In general, there were more looks to the reflexive's referent than towards the pronoun's referent for both adults and children in the cleft and in the no cleft condition, although the difference in looks to the two referents was more prominent in the reflexive condition. Interestingly, adults' attention to the reflexive referent is greater than to the pronoun referent from before the onset of the referential (0 ms) and decreases over time. Since the

36

reflexive referent is preceding the referential expression, it is not surprising that attention to this referent is greater even before the onset of the reflexive/pronoun. For the end of the segment, there is no clear preference for either referent in the cleft condition for pronouns (Figure2.6a left panel) as well as in the no cleft condition for reflexives (Figure 2.6b right panel). A preference can be observed for the pronoun referent in the no cleft condition (Figure 2.6a right panel), and for the reflexive referent in the cleft condition (Figure 2.6b left panel). For children on the other hand, looks to both referents are the same initially and start increasing for the reflexive's referent with the onset of the referential expression. This could indicate a delay in referent processing. Towards the end of the segment there is no clear preference for either referent for cleft and no cleft conditions with pronouns and for the no cleft condition with reflexives (Figure 2.6a, Figure 2.6b right panel). The cleft condition with reflexives reveals a clear preference for the reflexive referent (Figure 2.6b left panel).

The models for the second segment included reflexive referent advantage looks (elogits of reflexive referent looks - pronoun referent looks) as the dependent variable, the combined factor of condition (cleft, no cleft), *Referential Expression* (p= pronoun, r= reflexive), and *Group* (adults, children) as a **parametric term**, as well as smooth terms for the interaction between the combined predictor with *Time*. Further, the models included random intercepts for *Event* (a unique combination of *Subject* and *Trial*), and random smooths for *Subject* and *Item* to account for individual variation between participants. Table 2.3 shows the model summary. From the parametric coefficients, we can see that the model predicted significantly more looks to the reflexive's referent for adults in the cleft condition with reflexives than with pronouns (the intercept). It shows also significantly fewer looks by adults to the reflexive's referent for pronouns in no cleft sentences than with clefts. And lastly, it predicted significantly more looks to the reflexive's referent for reflexives in the no cleft condition than with clefts, for both adults and children.

For Segment 2, all smooth terms, except for combined factors *cleft_p_Adults* and *no cleft_r_Children*, were significantly different from 0 and non-linear (edf > 1)

To visualize the GAMM's predictions and understand how the effects con-

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.3915	0.2415	1.6210	0.1050
condCPGcleft_p_Children	-0.0765	0.4076	-0.1876	0.8512
condCPGcleft_r_Adults	1.3422	0.1682	7.9799	< 0.0001
condCPGcleft_r_Children	1.1665	0.4036	2.8903	0.0038
condCPGno cleft_p_Adults	-0.3772	0.1827	-2.0650	0.0389
condCPGno cleft_p_Children	-0.0738	0.3770	-0.1957	0.8449
condCPGno cleft_r_Adults	0.8168	0.1883	4.3381	< 0.0001
condCPGno cleft_r_Children	0.8803	0.3884	2.2663	0.0234
B. smooth terms	edf	Ref.df	F-value	p-value
s(Time):condCPGcleft_p_Adults	1.0131	1.0166	9.8326	0.0016
s(Time):condCPGcleft_p_Children	8.1040	8.6151	6.7199	< 0.0001
s(Time):condCPGcleft_r_Adults	6.9733	7.7467	87.7464	< 0.0001
s(Time):condCPGcleft_r_Children	5.4314	6.3415	6.9588	< 0.0001
s(Time):condCPGno cleft_p_Adults	7.5150	8.1668	10.3025	< 0.0001
s(Time):condCPGno cleft_p_Children	2.2423	2.6674	0.2100	0.8189
s(Time):condCPGno cleft_r_Adults	8.4801	8.7884	79.3366	< 0.0001
s(Time):condCPGno cleft_r_Children	2.1635	2.5729	0.6458	0.6185
s(Event)	1961.9186	2116.0000	29.0201	< 0.0001
s(Time,Subject)	547.3696	619.0000	403.6104	< 0.0001
s(Time,Item)	218.4983	305.0000	732.9118	< 0.0001

Table 2.3: Model summary for gaze data of second segment.

tribute to the participants' attention towards the possible referents and how they unfold over time, differences among these conditions were plotted. As above, we use the itsadug package to visualize the results (version 2.3, van Rij et al., 2020b). On the y-axis are the estimated differences between the looks to the reflexive's referent and looks to the pronoun's referent. A positive value indicates more looks towards the reflexive's referent, while a negative value indicates more looks towards the pronoun's referent in the condition first mentioned in the plot title. On the x-axis, 0 ms is the onset of the pronoun/reflexive.

Figure 2.7 shows the differences between adults and children in the cleft condition for stories containing pronouns (Figure 2.7a) and reflexives (Figure 2.7b). The dashed lines depict the significant differences between the comparisons. The first significant time window for pronouns (-200 - 84 ms) will not be taken into account since this time window starts before the onset of the pronoun. The other significant time window (3576 - 3820 ms) shows that adults looked more towards

Figure 2.7: Difference plots for Segment 2 comparing adults and children's looks to referents. Shaded areas represent 95% confidence interval. Significant time windows are marked with dashed lines.
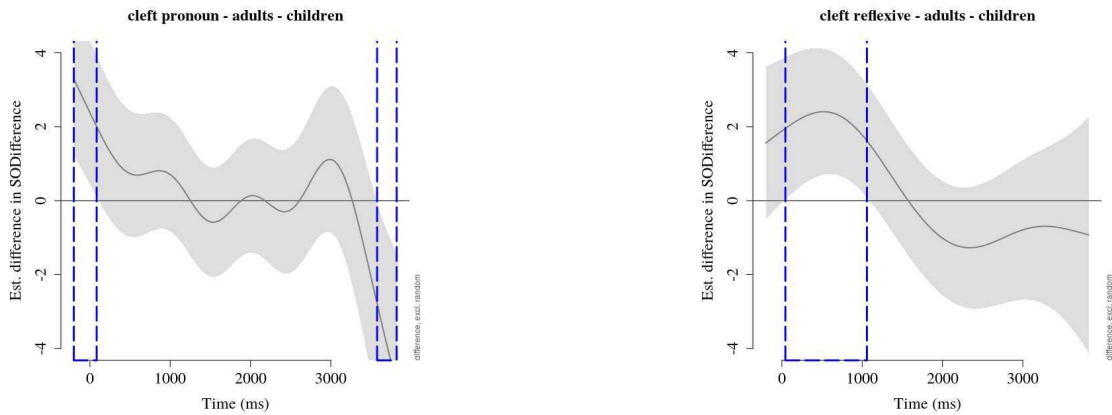
(a) Comparison of adults and children in cleft condition with pronoun.

(b) Comparison of adults and children in cleft condition with reflexives.

the pronoun referent than children. For the reflexives, the difference depicted (43 - 1058 ms) shows more looks to the reflexive's referent for adults than for children. For the no cleft conditions, there is no significant difference between adults and children with reflexives. For pronouns in no clefts, the significant time window (-200 - 124 ms) starts before the onset of the pronoun. Again, this window will be disregarded.

Figure 2.8 shows the differences between the comparisons of the various conditions for adults. Figure 2.8a depicts two significant time windows, one from 612 ms to 1424 ms, the other starting at 2602 ms to 3414 ms. Both windows depict more looks to the reflexive's referent for the cleft condition when compared to the no cleft condition for pronouns. For Figure 2.8b, there are also two significant time windows (328 - 1302 ms and 2439 - 3820 ms). Here as well, adults looked more to the reflexive referent with cleft than with no cleft conditions. The comparisons between the pronoun and the reflexive conditions for (Figure 2.8c) and for (Figure 2.8d) depict one significant time window each (206 - 3820 ms and 328 - 3454 ms). For both comparisons there more looks to the pronoun referent with pronouns than with reflexives.

Figure 2.9 shows the differences between the comparisons of the various condi-

Figure 2.8: Difference plots for Segment 2 comparison for adults. Shaded areas represent 95% confidence interval. Significant time windows are marked with dashed lines.
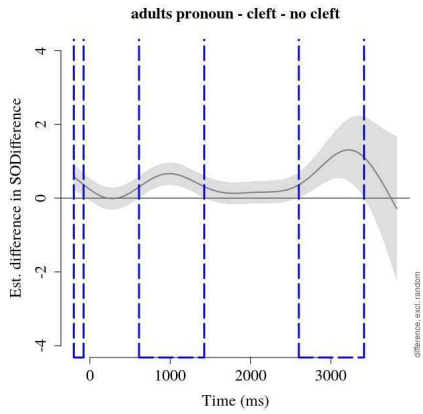
(a) Comparison of cleft and no cleft conditions with pronouns for adults.
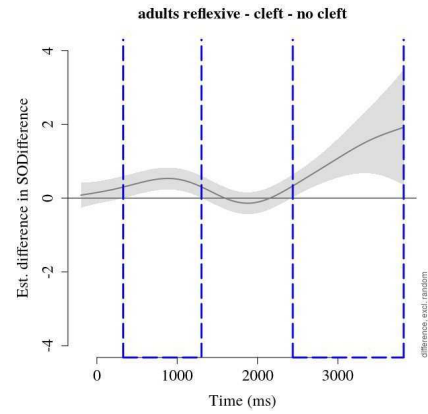
(b) Comparison of cleft and no cleft conditions with reflexives for adults.
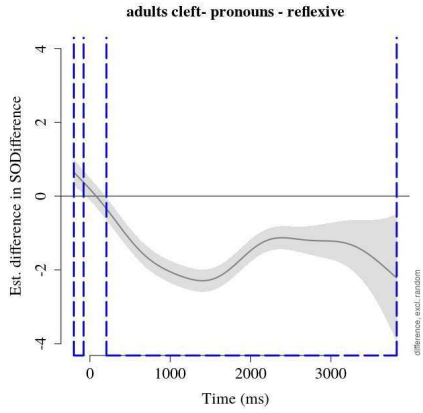
(c) Comparison of pronouns and reflexives in cleft conditions for adults.
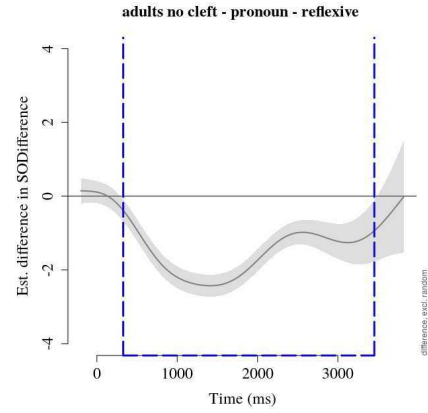
(d) Comparison of pronouns and reflexives in no cleft conditions for adults.

tions for children. Comparing cleft to no cleft condition for pronouns (Figure 2.9a), the first significant time window (856-1099 ms) shows more looks towards the pronoun referent in the cleft condition than in the no cleft condition. The second time window (3617 - 3820 ms) shows more looks to the reflexive referent with clefts than with no cleft sentences. Both windows are relatively short. For the comparison of reflexives in cleft and no cleft conditions (Figure 2.9b), the first significant time window (206-936 ms) shows more looks to the pronoun referent in the cleft condition than in the no cleft condition. The second significant time window (1870-2967 ms) shows more looks to the reflexive referent for the cleft condition than for the

Figure 2.9: Difference plots for Segment 2 comparison for children. Shaded areas represent 95% confidence interval. Significant time windows are marked with dashed lines.
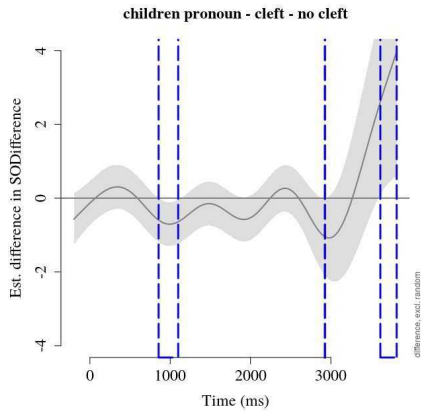
(a) Comparison of cleft and no cleft conditions with pronouns for children.
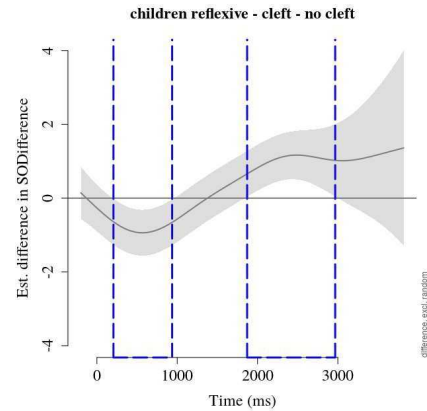
(b) Comparison of cleft and no cleft conditions with reflexives for children.
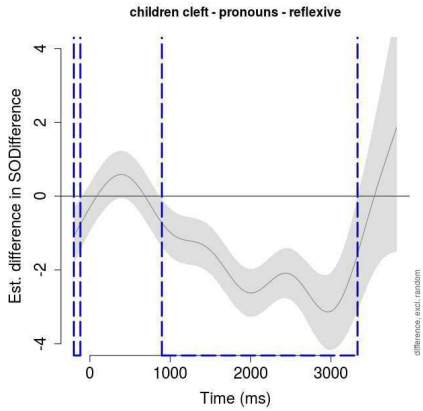
(c) Comparison of pronouns and reflexives in cleft conditions for children.
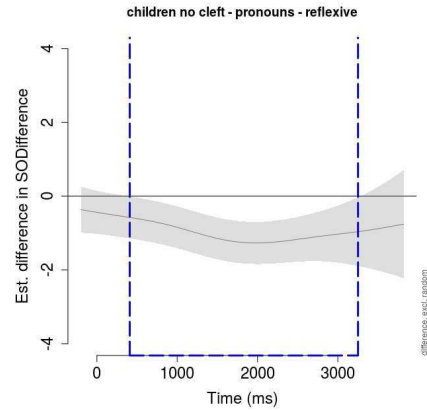
(d) Comparison of pronouns and reflexives in no cleft conditions for children.

no cleft condition. The comparison between pronouns and reflexives for the cleft (Figure 2.9c) and no cleft conditions (Figure 2.9d) show a similar pattern. There are more looks to the pronoun referent in pronoun condition than in reflexive condition for both cleft (896 - 3333 ms) and no cleft sentences (409 - 3252 ms).

To sum up the results, the comparison of adults and children in the no cleft condition revealed no differences in looks, neither for pronouns nor reflexives. For the cleft condition, there was only a significant difference with reflexives, where adults looked more to the reflexive's referent than children. For the comparison of cleft and no cleft conditions in adults, the gaze patterns were similar for pro-

nouns and reflexives. There were more looks to the reflexive's referent in the cleft condition than in the no cleft condition. The gaze patterns for the comparison of pronouns and reflexives in adults showed that with the cleft and with the no cleft conditions there were more looks to the pronoun's referent for the pronoun than for the reflexive. For children, the comparison of cleft and no cleft conditions for reflexives showed initially more looks to the pronoun's referent which later switched to more looks to the reflexive's referent for the cleft condition. For pronouns, there were only two short time windows that depicted significantly more looks to the pronoun's referent in the cleft condition than in the no cleft condition. For the comparison of pronouns and reflexives, children looked more to the pronoun's referent when hearing pronouns than with reflexives, both in cleft and in no cleft conditions.

Segment 3

The raw gaze data for Segment 3 show the proportion of looks towards the reflexive's referent and the pronoun's referent in the congruent and incongruent conditions with pronouns and reflexives for adults (Figure 2.10a) and children (Figure 2.10b).

(a) adults   *(b) children*
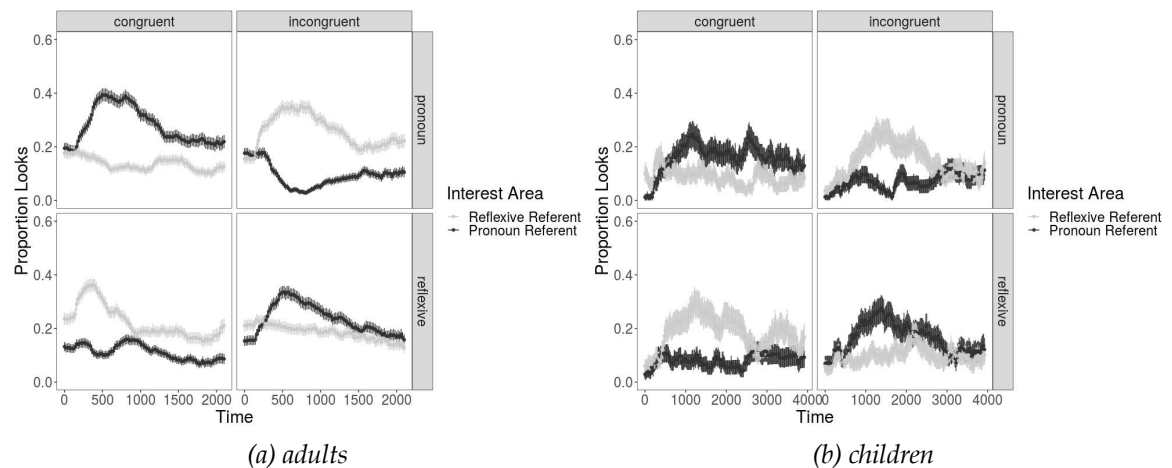
Figure 2.10: Average proportion of looks in Segment 3 in congruent and incongruent conditions.

For adults (Figure 2.10a), in the congruent condition attention to the target referents was higher than to the competitor over the course of the segment. Especially

in the beginning of the segment, there were more looks to the pronoun referent (target) after hearing a pronoun, and more looks to the reflexive referent (target) after hearing a reflexive. For the incongruent condition, attention to the competitors was higher than to the target referents for most of the segment. Initial proportion of looks to the reflexive referent (competitor) after hearing a pronoun, and initial proportion of looks to the pronoun referent (competitor) after hearing a reflexive were particularly high.

For children (Figure 2.10b), the initial looks to either referent did not reveal a clear preference in any condition. In the congruent condition later on there were more looks to the pronoun referent for pronouns, and more looks to the reflexive referent for reflexives. In the incongruent conditions looks to the reflexive referent increased over time for the pronoun, while with reflexives there were more looks to the pronoun referent.

In sum, the general gaze patterns for adults and children in Segment 3 were similar. The data revealed that participants paid more attention to the correct referent in the congruent condition while in the incongruent condition, participants were looking more at the incorrect competitor referent. The greater attention towards a certain referent could reflect a verification process during the truth value judgement task. This means that in the incongruent condition participants looked more towards the incorrect referent to verify that this referent indeed was not the target referent, whereas for the congruent condition they paid more attention to the correct referent to confirm that it was the target referent.

In the next step, we used GAMMs to analyze differences in the various conditions. The best fit model for this segment included reflexive referent advantage looks (elogits of looks to the reflexive referent - elogits of looks to the pronoun referent) as the dependent variable, as well as a combined factor of *Referential Expression* (pronoun, reflexive) *Congruency* (congruent, incongruent), and *Group* (adults, children) as a parametric term, as well as smooth terms for the interaction between the combined predictor with *Time*. Further, the models included random intercepts for *Event*, and random smooths for *Subject* and *Item* to account for individual variation between participants. We also accounted for autocorrelation by adding an

AR1 parameter to the model which specified the autocorrelation coefficient (*rho* = 0.4) and the starting point for each time series (Baayen et al., 2018). Note that the current model did not include *Condition* as a predictor since its effect was not significant.

Table 2.4 shows the model summary of the inferential statistics for the optimum-fit model for Segment 3. For the parametric terms (time course is not taken into account), the p-value reveals whether a combination of the combined factor was significantly different from the reference level *CombPredictp_c_Adult* (pronoun + congruent + adult) which is represented as the intercept. Positive *Estimate* values indicate more looks to the reflexive referent overall, while negative values signal more looks to the pronoun referent compared to the intercept. The model predicted *CombPredictp_i_Adult* (pronoun + incongruent + adult), and *CombPredictr_i_Adult* (reflexive + incongruent + adult) to be significant, for both there were more looks to the reflexive referent than for the intercept. Other combined factors of Condition and Group were not significantly different from the intercept. For Segment 3, all smooth terms, except for the the combined factor *CombPredictr_c_Child* conditions, were significantly different from 0 and non-linear.

For the visualisations of the GAMM's predictions and to understand how the effects might contribute to the participants' attention towards the possible referents and how they unfold over time, we again used the *itsadug* package (version 2.3, van Rij et al., 2020b). On the y-axis are the estimated differences between the looks to the reflexive's referent and looks to the pronoun's referent. A positive value indicates more looks towards the reflexive's referent, while a negative value indicates more looks towards the pronoun's referent. On the x-axis, 0 ms is when the new picture was displayed (see Figure 2.2b for an example). Significant differences in looks to the two target referents are marked with dashed lines.

For adults, the first significant time window was found between the incongruent and congruent conditions when participants previously heard sentence containing a pronoun (Figure 2.11a). In the significant time window (0 - 2252 ms) there were more looks to the reflexive's referent for the incongruent condition than for the congruent condition. Since the target referent for pronouns is the pronoun

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-0.5183	0.4940	-1.0493	0.2940
CombPredictp-c-Child	0.1779	0.1959	0.9079	0.3640
CombPredictp-i-Adult	1.6435	0.6986	2.3526	0.0186
CombPredictp-i-Child	1.0535	0.7106	1.4825	0.1382
CombPredictr-c-Adult	-7.8689	12.7098	-0.6191	0.5358
CombPredictr-c-Child	0.9820	0.7103	1.3825	0.1668
CombPredictr-i-Adult	0.4225	0.1211	3.4892	0.0005
CombPredictr-i-Child	-0.1044	0.2228	-0.4686	0.6394
B. smooth terms	edf	Ref.df	F-value	p-value
s(Time):CombPredictp-c-Adult	6.6071	6.7748	4.1441	0.0022
s(Time):CombPredictp-c-Child	7.3054	7.4225	2.5397	0.0258
s(Time):CombPredictp-i-Adult	5.6435	5.8150	2.2868	0.0410
s(Time):CombPredictp-i-Child	4.8353	5.2335	7.7277	< 0.0001
s(Time):CombPredictr-c-Adult	7.9150	7.9966	49.5805	< 0.0001
s(Time):CombPredictr-c-Child	1.0006	1.0006	0.6446	0.4221
s(Time):CombPredictr-i-Adult	7.9818	8.1584	5.0094	< 0.0001
s(Time):CombPredictr-i-Child	5.6682	5.8732	8.4254	< 0.0001
s(Event)	1825.7182	1941.0000	47.9168	< 0.0001
s(Time,Subject)	323.0357	590.0000	364.3359	< 0.0001
s(Time,Item)	165.9615	308.0000	4695.9998	< 0.0001

Table 2.4: Model summary for gaze data of third segment.

referent, this means there were more looks to the competitor referent in the incongruent condition than in the congruent condition. The next significant time window (0 - 477 ms) showed more looks to the pronoun's referent for *congruent + pronoun* than for *congruent + reflexive* (Figure 2.11b).

For the comparison between the conditions *congruent + pronoun* and *incrongruent + reflexive* (Figure 2.11c), there were more looks towards the pronoun's referent for *congruent + pronoun* than for *incrongruent + reflexive* during the entire segment. The second comparison between *incongruent + pronoun* and *congruent + reflexive* (Figure 2.11d) showed two significant time windows. Again, in the first significant time window (0 - 185 ms) there were more looks to the pronoun referent for *incongruent + pronoun* than for *congruent + reflexive*. The second time window (477 - 2226 ms) showed more looks to the reflexive referent for *incongruent + pronoun* than for *congruent + reflexive*.

For children, the first significant time window (1749 - 2252 ms) was for the

Figure 2.11: *Difference plots for Segment 3 comparisons for adults. Shaded areas represent 95% confidence interval. Significant time windows are marked with dashed lines.*
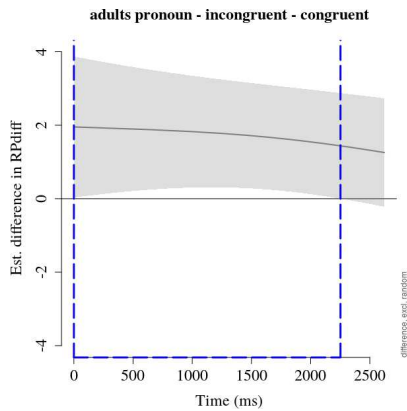

(a) Comparison of incongruent and congruent conditions with pronouns for adults.
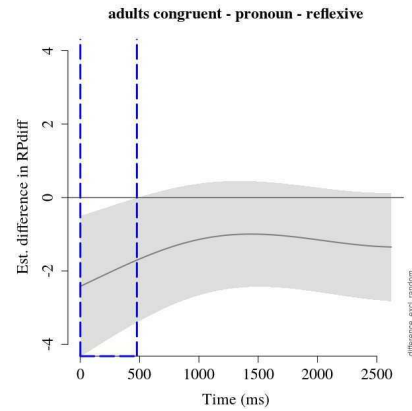
(b) Comparison of pronouns and reflexives in congruent condition for adults.
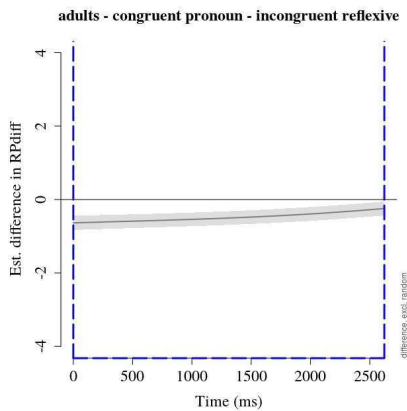
(c) Comparison of congruent + pronoun and incrongruent + reflexive for adults.
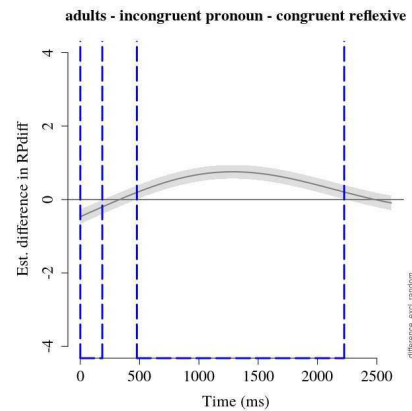
(d) Comparison of incongruent + pronoun and congruent + reflexive for adults.

comparison between the congruent and incongruent conditions with pronouns (Figure 2.12a). Children looked more to the reflexive referent in the incongruent condition than in the congruent condition. This means the competitor referent was more attended in the incongruent than in the congruent condition. The other significant time window (0 - 556 ms) was found when comparing *incongruent + pronoun* to *congruent + reflexive* (Figure 2.12b). Here, children looked more to the reflexive referent with the *congruent + pronoun* condition than with the *incongruent + reflexive* condition.

When directly comparing adults and children's gaze patterns, the only signifi-

Figure 2.12: Difference plots for Segment 3 comparisons for children. Shaded areas represent 95% confidence interval. Significant time windows are marked with dashed lines.
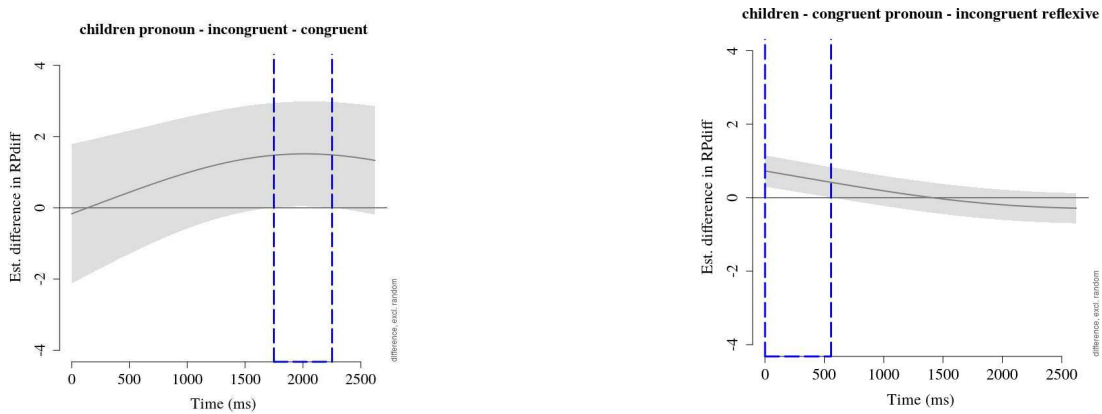
(a) Comparison of incongruent and congruent conditions with pronouns for children.

(b) Comparison of congruent pronouns and incongruent reflexives conditions for children.

cant difference in looks was present when comparing *adults + pronoun + incongruent* to *children + pronoun + congruent* (Figure 2.13a). For adults, there were more looks to the reflexive referent than for children (1387 - 2278 ms). This difference is not surprising since the two compared conditions were found to have different referent preferences. It is more surprising that other comparisons between conditions with different referent preferences were not significantly different in their gaze pattern. This may be due to children's gaze patterns being flatter than that of adults (peaks are lower). Moreover, we can also observe a delay for increases in looks to a certain referent. Another smaller significant time window (1722 - 1881 ms) was found for the comparison between *adults + reflexive + incongruent* and *children + pronoun + congruent* (Figure 2.13b). There were more looks to the reflexive for adults than for children. This difference is most likely due to a second peak in the child data, whereas for adults the looks to the referent were decreasing. However, having no difference in looks for the conditions in which referent preferences align also confirmed that the general gaze pattern for adults and children was similar.

In summary, the gaze pattern for the third segment is generally similar between adults and children: more looks to the correct referent with congruent conditions, more looks to the incorrect referent with incongruent conditions. However, while

Figure 2.13: Difference plots for Segment 3 comparisons between adults and children. Shaded areas represent 95% confidence interval. Significant time windows are marked with dashed lines.
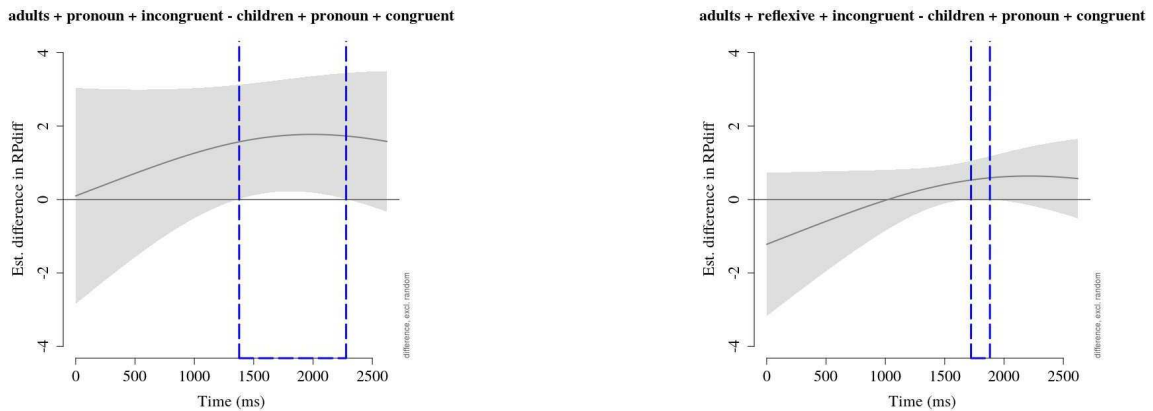
(a) Comparison of adults incongruent pronoun and children congruent pronoun conditions.

(b) Comparison of adults incongruent reflexive and children congruent pronoun conditions.

the general gaze pattern followed the same tendency there were still differences between adults and children. First, the comparison between pronouns and reflexives for congruent conditions was only significant for adults. Next, the comparison for the *congruent pronoun* to *incongruent reflexive* was only significantly different for adults. Children's gaze pattern is "flatter" than that of adults'. This means that children paid overall less attention to the two target referents and were more distracted by the entire picture than adults. This might be why some differences were only visible in the adults' data. Moreover, less attention to the target referents in turn may have also affected children's performance on the truth value judgement task and led to more errors.

Interestingly, for adults there was more attention towards the reflexive referent in *incongruent + pronoun* than in *congruent + reflexive* and also more attention towards the pronoun referent in *congruent + pronoun* than in *incongruent + reflexive*. For this difference incongruency by itself cannot be the explaining factor. If it were only for incongruency then *incongruent + reflexive* should have more looks to the pronoun referent than *congruent + pronoun* which was not the case. The common factor for both comparisons was the pronoun. So, when a pronoun was involved, attention towards the referent was greater than when a reflexive was in-

volved. This could hint at differences in processing for pronouns and reflexives. However, this hypothesis cannot be completely confirmed for children, as only the comparison between *incongruent + pronoun* and *congruent + reflexive* showed significant differences in the gaze pattern. Here too, children looked more towards the reflexive referent with the pronoun than with the reflexive.

2.5 Discussion and Conclusion

In this study, we investigated the effects of *it*-clefts on the processing of reflexives and pronouns in German children and adults in order to assess whether processing of the preceding discourse context resulted in differences of reference processing and whether this would help explain why children's pronoun interpretation in some languages was not yet adult-like.

The results of Segment 1 revealed that after an initial decrease in looks to the subject referent, the cleft indeed drew more attention to the subject (clefted) referent. This was the case for both adults and children. However, we also found differences in the gaze behaviour between children and adults in the cleft condition. Children's peaks in looks appeared later than adults'. Instead of children just being generally slower as has been previously found (e.g. Contemori & Marinis, 2014; Hartshorne, Nappa & Snedeker, 2015; Thompson, Driscoll & Markson, 1998), the observed difference might actually point towards a delay in processing caused by the cleft (see below for further discussion), since the general gaze pattern for children and adults seemed to be the same and there were no differences for the no cleft conditions.

In contrast to the gaze data from Segment 1, the results of the Truth Value Judgement task showed that clefting the pronoun's referent did not affect the final interpretation of either pronouns or reflexives. This is in line with Blything et al.'s (2021b) findings and supports their account of focusing only making the referent more distinct in working memory (cf. Birch et al., 2000; Foraker & McElree, 2007), but not more likely to become the pronoun's referent. It is also not surprising considering that the contexts for the pronominal forms in this study were

49

not ambiguous and each form had its intended referent.

For Segment 2, the gaze data suggest that children's processing of pronouns and reflexives is adult-like if the sentence follows a no cleft structure, as there was no difference between adults and children with regards to the attention towards the referents for the reflexive and the pronoun. Also, the overall pattern with regards to referent attention was similar: more looks to the reflexive's referent with reflexives and conversely more looks to the pronoun's referent when the pronoun was the target referential expression.

The were no significant differences in the gaze pattern for adults and children in the cleft condition when a pronoun was present. However, for reflexives, adults looked significantly more to the reflexive's referent than children. The difference could possibly be due to the cleft affecting attention towards the two referents differently than in no cleft sentences. While adults overall paid more attention to the pronoun referent with pronouns than with reflexives, the presence of the cleft shifted attention to the reflexive referent. The effect was not limited to reflexives but was also observed for pronouns, since looks to the pronoun referent were significantly reduced compared to the no cleft condition. We had assumed that the cleft would draw more attention to the pronoun's referent which did not appear to be the case for adults. These findings suggest that adults' processing of the cleft affected attention to referents differently when compared to children. A possible explanation for these results could be that the use of the cleft did not result in narrow focus, i.e. focusing the clefted referent only, but rather in broad/dual focus, i.e. focusing both referents of the main clause. This means that clefts were interpreted as broad focus (Frascarelli & Ramaglia, 2013; Hedberg, 2013; Krifka, 2008; Lambrecht, 2001; Prince, 1978; Tönnis, 2022, for different possibilities for the information structure of clefts see e.g.). Since the first sentence introduced four referents, it could be that clefting focused the referents that remained in the discourse, therefore increasing the attention towards both referents. Generally, clefts resulted in more attention towards the reflexive referent for adults. For reflexives, clefting led to more looks towards the correct referent, whereas for pronouns, clefting led adults to look less often at the correct referent.

For children, the presence of clefts seemed to have a different effect on subsequent processing. Their gaze data suggests that the pronoun referent indeed received more attention in the cleft condition than in the no cleft condition for both pronouns and reflexives. The attention toward the pronoun's referent right after the onset of the referential expression was even more prominent in the reflexive condition when compared to the pronoun condition. A possible explanation could be that if the cleft indeed increased the attention to the pronoun's referent, competition between the two referents would be greater in the cleft condition compared to the no cleft condition. The additional attention towards the pronoun's referent for the reflexive seemed to cause competition, changing attention to the incorrect referent at first before shifting attention to the correct referent.

This interpretation of the data indicates that information structure in the form of cleft structures affected processing of pronouns and reflexives in different ways for adults and children. This in turn could mean that clefts were interpreted differently. For adults, the cleft might have led to a broad/dual focus interpretation, while for children it appeared to only focus the clefted referent. This could mean that the context preceding the cleft also influenced how the cleft effect unfolded over the remaining discourse. As mentioned above, as the first sentence introduced four referents, adults might have interpreted the cleft as signaling broad/dual focus in the sense of "focus on the remaining two referents, disregard the other two". If that was the default interpretation for this cleft-structure, then, interestingly, this suggests that unlike adults, children might not take the global discourse context into account, as their cleft processing seemed to have led to a narrow focus interpretation, but instead were processing reference more locally.

As mentioned earlier, the raw gaze data might indicate a delay in processing of referents which could also partly explain the differences between children and adults. Similarly, Sekerina et al. (2004) have found children's attention to shift about 1000 ms later than adults' (see also Hartshorne et al., 2015). In contrast, van Rij et al. (2016) as well as Clackson et al. (2011) found children's gaze behavior not to be much slower than that of adults. While these findings only apply to the gaze behavior after the onset of the pronoun and do not reflect what was happening

before the onset of the pronoun, our gaze data of Segment 1 also points towards a delay in processing. Even when we observe some differences, the general gaze patterns for adults and children were not fundamentally different with regards to general tendencies for the referents. So, a delay in processing may also be an explanation as to why gaze patterns in Segment 2 differed between children and adults, potentially indicating further processing difficulties for children.

Lastly, another explanation for these differences in processing/attention between children and adults might be found in the addition of the relative clause in the cleft sentences. Compared to the no cleft sentence, cleft sentences contained a relative pronoun (*der*) and the (finite) verb placement changed from second (V2) to the final position. Studies investigating children's comprehension of relative clauses in German found that with subject relative clauses, young children performed above chance level (Diessel & Tomasello, 2005; Kidd, Brandt, Lieven & Tomasello, 2007, for 3- and 4-year olds) while older children perform at ceiling level (Arosio, Yatsushiro, Forgiarini & Guasti, 2012, for 7-year olds). Moreover, Aravind, Hackl & Wexler (2018) found that children did not experience difficulties comprehending *it*-clefts if these were presented in felicitous contexts. But if the cleft structure indeed affected the processing of the referential expression, it seems odd that this effect was more prominent in children (with respect to the length of significant time window) for reflexives (see Figure 2.9b) than for pronouns (see Figure 2.9a), since reflexives have been claimed to impose no comprehension difficulties on children.

To sum up, the differences between adults and children found in the gaze data can be interpreted in two main ways: (i) children are simply slower at processing than adults which leads to seemingly different gaze patterns between the two groups; (ii) children and adults process clefts differently which in turn affects their processing of pronouns and reflexives.

The first explanation (i) describes the differences found in children and adults in terms of a general processing delay in children. Children take longer to integrate the information of the discourse during sentence processing which is then reflected in their slower gaze pattern (see also Sekerina et al., 2004). This explanation seems

probable for Segment 1 (compare Figure 2.5a and Figure 2.5b) as the gaze patterns align in terms of looks to referents and the significant time windows also point out the same difference in adults and in children, whereas Segment 2 (compare Figure 2.8a and Figure 2.9a, and Figure 2.8b and Figure 2.9b) shows different preferences in looks to referents. A delay in processing does not seem to completely capture the differences found in the gaze pattern between adults and children, nevertheless slower processing in children may still contribute to it.

Turning to explanation (ii), children processing the preceding context differently affects the pronoun processing as well. This idea would be in line with the findings from van Rij et al. (2016). In their study, when the preceding discourse was limited to one referent, children's processing was more adult-like than processing of two-referent contexts: no differences were found in the gaze data between children and adults in one referent context, whereas this was not the case for two-referent contexts. The more adult-like processing pattern also affected referential choice. Children's performance on the pronoun interpretation task was more adult-like than in discourse contexts with two referents.

However, it does not explain the cross-linguistic difference found in children. Why would Dutch, but not German children have difficulties interpreting pronouns, if both child groups' processing differs from that of adults? So, the processing difference by itself cannot be the explanation for children's performance. On a linguistic level what is different between German and Dutch is the function of their pronouns, as discussed in the introduction. The German pronominal system shows functionally more complementarity than Dutch, it does not allow pronouns to be linked to the local subject referent, while in Dutch this is possible in certain contexts (see example 5). The Dutch pronoun seems more referentially ambiguous than the German pronoun and as Sekerina et al. (2004) have shown, children are indeed aware of the referential ambiguity. This means, it does not matter if German children cannot take discourse/pragmatic information into account in the same way as adults, since the pronoun systems constraints referential links, whereas if Dutch children cannot take discourse/pragmatic information into account, it can lead them to the incorrect referent, because their pronoun system is referentially

not as restricted as the German one. Therefore, shifting attention to the intended referent may be beneficial when processing the pronoun, but for interpretation in German it is not needed since the pronoun system is restricting referential links to incorrect referents, which is not the case in Dutch. Interpreting pronouns in Dutch requires discourse/pragmatic information, which children do not take into account due to limited working memory (also cf. Ruigendijk et al., 2010). Instead, they use the local discourse context (see also Trueswell, Sekerina, Hill & Logrip, 1999, for children's processing of temporarily ambiguous sentences). In the local context, they pay more attention to the clefted referent as can be seen in children's gaze data of Segment 1. The cleft makes the referent more distinct in working memory – as reflected in children's gaze behaviour in Segment 2 – which in turn makes it easier to retrieve. Therefore, in contrast to adults, children may depend to a greater extent on discourse information rather than including pragmatic information about the pronoun (e.g. Rule I) which would explain why in Spenader et al. (2009) and van Rij et al. (2016) children's performance improved when the pronoun's preceding context was modulated.

Thus, modulations in the discourse context with respect to the possible referents may facilitate children linking the pronoun to the correct referent as it may draw children's attention away from the incorrect referent. The reason for children to rely more on discourse instead of pragmatic information to establish referential links for the pronoun may be because it is easier on their limited working memo- ry capacity. The information in the discourse is directing attention to a specific referent, so the "memory boost" the referent would receive from being focused is easier to access than pragmatic knowledge about the pronoun. The pragmatic knowledge would require children to evaluate whether this specific context allows the pronoun to be resolved towards a certain referent (see also Ruigendijk et al., 2010). So, the discourse course information would be more readily available while pragmatic knowledge would still need to be computed when encountering the pronoun on top of processing discourse information.

Another note on the idea of the information structure effect unfolding differently in the succeeding discourse for children and adults, and it affecting attention

to referents in different ways: Similar to Järvikivi et al.'s study (2014), children's gaze pattern deviated from adults' in that there was more attention to the clefted referent. However, in our study even adults were affected by the presence of the cleft. The difference in cleft effects may be due to differences in the context in which the clefts were used. While Järvikivi et al. (2014) only introduced two referents in the preceding discourse, in this study we introduced four referents. In contrast, Blything et al.'s study (2021b) also introduced four animal characters in the first sentence of their stimuli, but additionally the cleft sentence was preceded by a subject question. Children's gaze pattern followed adults' more closely in that the subject cleft condition, which was preceded by a subject focus question, also led to more looks to the subject referent when compared to a baseline broad focus condition. However, since the authors analyzed adult and child data separately, there was no direct comparison of adults' and children's gaze pattern. Taken together, the results of all three studies suggest that the cleft effect can be modulated depending on the contexts preceding the cleft. Moreover, while the use of the cleft led to children and adults being shifted to different referents in the subsequent discourse, the effect was similar for pronouns and reflexives. For adults, the reflexive referent was fixated more than the pronoun referent with both reflexives and pronouns, whereas children paid more attention to the pronoun referent than the reflexive referent with reflexives and pronouns. This may suggest that for adults too pronoun and reflexive processing is guided by discourse information first before other linguistic factors are taken into account (cf. Clackson et al., 2011). For Segment 3, we found the general gaze pattern for children and adults to be similar for the congruent and the incrongruent conditions as well as for reflexives and pronouns. The gaze data showed more attention was paid to the correct referent in the congruent conditions which was also reflected in the final interpretation of the pronoun/reflexive sentence. In this case, looks to a certain referent were predicting the final interpretation of the pronoun/reflexive. However, in the incongruent conditions the participants looked more towards the incorrect referent, i.e. when the previous sentence contained a reflexive, they paid more attention to the pronoun's referent and vice versa. For adults, accuracy of their final interpretation

was also slightly, but significantly, lower compared to the congruent conditions. For children, accuracy was far lower in the incongruent conditions. Children's performance can be (partially) explained by a "yes"-bias (Chien & Wexler, 1990, cf.), as they seemed to prefer visual information, but possibly earlier attention to the incorrect referent may have influenced final interpretation, as well. In combination with uncertainty of whether the picture displayed a congruent or incongruent scene, attention towards the incorrect referent may have led participants to accept the incorrect interpretation. This means that participants, especially children, relied more on the visual cue for interpretation when visual and linguistic information did not align. Another potential reason for the low performance in the incongruent condition could be that the picture did not provide sufficient visual information as the action of the verb was not visualized during but only after sentence processing. In addition, overall attention to the referents seemed lower for children than for adults as peaks of proportions of looks were lower with children compared to adults. This finding may also explain why even in the congruent conditions children's accuracy was lower than adults' as it seems to imply that even when linguistic and visual information align, lower attention may lead to more mistakes during the interpretation process. Thus, attention towards a certain referent cannot simply be taken as reflecting participants linking the pronoun or reflexive to that referent. Yet we cannot completely dismiss the influence attention may have on the interpretation of reflexives and pronouns. While it did not seem to have a major influence on interpretation, it still seemed to affect participants' choices to some degree, especially when input cues did not align (Blything et al., 2021a,b; Clackson et al., 2011).

Another interesting finding from Segment 3 points towards possible differences in processing of reflexives and pronouns in adults. Generally, gaze patterns for congruent conditions were mirrored in the incongruent conditions, however, when running a comparison between the conditions that preferred the same referent (e.g. in *congruent + pronoun* and *incongruent + reflexive* conditions participants fixated the pronoun referent more than the reflexive referent), attention towards referents was higher when the previous sentence contained a pronoun. While reflexive interpre-

tation requires "only" structural processes, pronoun interpretation often times also involves non-syntactic factors like discourse contexts, making pronoun interpretation a more effortful process. Therefore, one explanation as to why adults paid more attention to "preferred" referents after hearing a sentence containing a pronoun than after sentences containing a reflexive may be that the greater attention reflects the more effortful process of pronoun interpretation. For children, this was only found in the comparisons for the conditions where the reflexive referent was the preferred referent. This too may indicate that pronouns are more effortful to process than reflexives.

As discussed in the introduction, studies looking into adults' processing of reflexives and pronouns did find evidence for pronouns requiring more processing efforts than reflexives. However, this was only the case for Dutch adults – a language where children were found to experience difficulties interpreting pronouns (Hendriks et al., 2011; Vogelzang et al., 2016). For German (Vogelzang et al., 2021b), there was no evidence for additional effort in pronoun processing. However, our results do not contradict these findings as this difference was only found in the last segment and not during the actual pronoun segment. The reason why we nonetheless assume that our findings can be linked to processing difference is because in our experimental design the picture was changed after participants heard the sentence. Only then could they compare the visual to the auditory stimuli and make a decision of whether or not the picture was congruent with the sentence.

The higher effort for pronouns could also reflect their ambiguity. In cross-linguistic explanation given above, it was said that German pronouns are stricter in their referential function. However, that does not mean that German pronouns are not ambiguous. The pronoun could also refer to an external referent (e.g. one of the other animal characters mentioned in the first sentence of the mini stories). So, in Dutch the referential function of the pronoun is extended and therefore adds more possible referents compared to German. In other words, the degree of ambiguity of the pronoun is higher in Dutch than in German. Children have more options to choose from and fail to reject the local referent, because their pragmatic knowledge is not completely established yet. Therefore, since children seem to

pay only attention to the immediate (local) linguistic context, shifting attention to the pronoun's intended referent may aid children in rejecting the incorrect referent. Our results show that attention to referents may affect final referent choice, as even adults showed slightly worse performance in the incongruent conditions in which the incorrect referent had received more attention.

Further, this paper made a methodological contribution. It showed that the use as well as analysis of eye-tacking data should not only focus on the pronoun segment, but rather should be extended. Including the segments preceding and following the pronoun segment allowed us to find differences in children's and adults' processing of cleft structures which may explain the difference in gaze data found in the pronoun segment. In addition, the final segment points out differences in pronouns and reflexives' gaze pattern in adults which may indicate higher processing efforts for pronouns. However, while some of these differences also occur in the child data, this hypothesis cannot be fully confirmed yet.

In conclusion, this study showed that children are sensitive to changes in information structure, as focusing via clefting indeed draws more attention to the clefted referent for both adults and children. However, processing of these factors is not yet adult-like leading to difference in attention towards referents during processing. The presence of the cleft led adults to attend the reflexive referent more for both reflexives and pronouns, whereas children fixated the pronoun referent more with reflexives and pronouns. So, while the information structure effect of the cleft unfolded differently for children and adults in the subsequent discourse, it affected pronouns as well as reflexives in a similar way, indicating that adults as well are affected by the local discourse when processing reflexives and pronouns. Our findings emphasizes the importance of discourse context for children's processing of pronouns and may explain why children experience pronoun interpretation problems in some languages. When information structure clearly directs attention to one referent, this information is more easily available to children than evaluating pragmatic contexts in order to exclude referents from being linked to the pronoun. This may also be where the cross-linguistic difference lies: While Dutch children would have to go through the evaluation process to correctly resolve the pronoun,

German children do not need to evaluate contexts since the pronoun is restricted in its referential function. With respect to limited working memory capacity, it may therefore be easier for Dutch children to rely more on information from the discourse than employ pragmatic knowledge.

Chapter 3

Nothing Else Matters – *He* likes Subject Referents

Abstract

I report the results of one visual-world eye-tracking experiment and two referent selection tasks in which we investigated the effects of information structure in the form of prosody and word order manipulation on the processing of subject pronouns *er* and *der* in German. Factors such as subjecthood, focus and topicality, as well as order of mention have been linked to an increased probability of certain referents being selected as the pronoun's antecedent and described as increasing this referent's prominence, salience or accessibility. The goal of this study was to find out whether pronoun processing is primarily guided by linguistic factors (e.g. grammatical role) or non-linguistic factors (e.g. first-mention), and whether pronoun interpretation can be described in terms of referents' 'prominence' / 'accessibility' / 'salience'. The results showed an overall subject preference for *er*, whereas *der* was affected by the object role and focus marking. While focus-marking affected attention towards the referents, this effect did not trans- fer to the final interpretation of *er*, suggesting that 'salience' does not contribute to referent selection. Overall, the results suggest a primacy of linguistic factors in determining pronoun resolution.

3.1 Introduction

In pronoun processing research, much of the research interest lies in finding out how listeners identify the referent of a pronoun when the previous discourse offers more than one possible referent. In general, pronouns are used when their intended referents appear to be recoverable from the preceding context (Givón, 1983; Gundel et al., 1993; Schwarzschild, 1999). In various theories of pronoun resolution, the degree of this recoverability has been captured using various other concepts, such as *accessibility* (Ariel, 1990, 2001; Arnold, 2010) or *most highly ranked element* (Gordon, Grosz & Gilliom, 1993; Grosz, Weinstein & Joshi, 1995). What most theories have in common is the assumption that several factors contribute to how 'prominent', 'salient', or 'accessible' a referent is, i.e. how easy or likely it is (in processing terms) that a referential link is established. However, while there is no agreement on how to define 'accessibility', 'prominence' or 'salience' (as stated e.g. for prominence by Ladd & Arvaniti, 2023), most definitions are circular as the properties that cause salience of a feature are the ones that follow from it being salient (see Boswijk & Coler, 2020, for an overview on salience). With respect to pronoun resolution, that means that factors that lead to a pronoun referent being chosen more often (such as it being a subject or first mentioned, see below) are said to make this referent more accessible or prominent because pronouns are assumed to refer to the most prominent referent (see also Kaiser & Trueswell, 2008, for a critique of this circularity). Moreover, it is not universally agreed on which factor(s) decide pronoun reference, whether or not mediated via prominence/accessibility. Factors such as syntactic (grammatical role and linear order) and thematic prominence (semantic roles), and information structure (topichood, focus) have been suggested to increase attention to referents and raise the predictability for a certain referent to occur again or increase the availability of these referents in memory (Arnold, 2010) (see also Arnold & Griffin, 2007; Arnold, Kaiser, Kahn & Kim, 2013; Blything et al., 2021b; Ferreira, Slevc & Rogers, 2005; Foraker & McElree, 2007; Gundel et al., 1993; Kehler, 2002).

In this paper, we aim to contribute to the clarification of both of these issues on

the basis of experimental data. As for prominence / accessibility, we will argue that theories of pronoun resolution centered on these notions are problematic, as they do not provide independent definitions of these concepts. Our experiments test a promising way of substantiating them, namely by connecting them to well-defined information structural notions, but show that neither focus nor topic or givenness reliably predict pronoun referent choice, though they guide participants' attention to referents. As for the factors that do guide the reference resolution, our experiment was designed to disentangle linguistic (subjecthood) and more general cognitive factors (first mention) which have been the subject of a long-standing debate, while also freeing them from information structural confounds. we find that two German subject pronouns differ in preferences, in line with Kaiser & Trueswell's (2008) form-specific approach to reference resolution, but are both guided by linguistic factors.

Regarding the first issue, the notions of prominence and accessibility have been described in various (not mutually exclusive) approaches in the literature on pronoun resolution. For instance, on the basis of Himmelmann & Primus' (2015) framework on prominence, von Heusinger & Schumacher (2019) argue that prominence in discourse can be identified as relational, dynamic and attracting linguistic operations. According to von Heusinger & Schumacher, prominence is a relational property that elects one element from a set of elements of equal type. With respect to pronoun resolution this means singling out a certain referent from a set of relevant referents. This step implies a ranking within the set of relevant referents which is argued to rely on so-called prominence-lending cues, e.g. grammatical role, thematic role, information structure role, or position. Further, von Heusinger & Schumacher state that prominence is dynamic as a referent's prominence depends on prominence-lending cues and these may change as the discourse unfolds; the same referent can be prominent at one point and be overtaken by another referent at a different point within the span of the discourse. Topic discontinuity (Givón, 1983) is listed as an example for the change in a referent's prominence and the high likelihood of prominent referents to be rementioned in discourse as an example of structural attraction.

A related proposal – the *Expectancy Hypothesis* (see e.g. Arnold, 2001; Arnold et al., 2007; Arnold & Griffin, 2007) – is that discourse cues, such as subjecthood, topicality or first-mention, correlate with the likelihood of a referent being mentioned again in the upcoming discourse: Accessible referents possess relatively high expectancy. The basis for the link between accessibility and expectancy is evidence that accessible or "prominently mentioned" entities are more likely to be subsequently mentioned (Arnold, 2010, p. 192).

Another approach relates salience/prominence/accessibility to attention (e.g. within Centering Theory see Brennan, 1995; Grosz & Sidner, 1986; Grosz et al., 1995). Based on the *Givenness Hierarchy* (Gundel et al., 1993) which links six cognitive statuses to the assumed attention state of listeners and the corresponding typical use of referential expressions (e.g. full noun phrase vs. demonstrative vs. pronoun), accessible referents receive more attention (e.g. Brennan, 1995). In turn, more attention likely leads to an enhanced representation of that referent in the discourse model (Foraker & McElree, 2007).

As pointed out earlier, however, defining prominence/accessibility in this way is problematic because accessibility is the explanation for why a pronoun can be used to refer to a certain referent, but at the same time the use of pronouns explains what accessibility is (cf. Arnold, 2010).

I also do not know what exactly makes one property, like subjecthood, more prominent than another, like objecthood. Moreover, with respect to the first two approaches, the question arises how we can test whether the chosen referent was indeed more prominent or accessible than any other given referents. we could only assume this to be the case based on which referent had been selected. However, if prominence / accessibility is to be an explanation for the interpretation of a pronoun, its referent has to be prominent during pronoun processing and even before the mention of the pronoun. One possible way to substantiate the notions of prominence and accessibility and to identify prominent / accessible referents independently of the final referent choice reported by listeners is to use a time-sensitive tool to assess attention towards the different referents within the discourse.[1]

---

[1]Note that while Gundel et al. (1993), Grosz & Sidner (1986), Grosz et al. (1995), Brennan (1995)

In our study we therefore applied eye-tracking as a tool to measure attention towards the critical referents during the time course of the sentence containing the ambiguous pronoun as well as the preceding sentence containing the referents.

Regarding the second issue, referentially ambiguous subject pronouns are often assumed to typically choose the subject of the preceding sentence as their antecedent, but it has long been debated whether this is best characterized as a subject (Crawley & Stevenson, 1990; Frederiksen, 1981; Gordon & Chan, 1995) or a first-mention preference (Gernsbacher, 1985; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989). For example, in (8), *der Schauspieler* 'the actor', the preferred referent for the pronoun *er* 'he', is both the subject and the first-mentioned referent in the preceding sentence. Thus, the underlying question is whether and to what extent it is linguistic structure or more general non-linguistic cognitive processes that drive the processing of subject pronouns.

(8)   Der Schauspieler hat den Koch angerufen, und zwar mit einem Handy. *Er* war zu diesem Zeitpunkt schon ziemlich m ü d e .
       'The actor called the cook, namely with a mobile phone. *He* was already pretty tired at this point.'

Grammatical role information, more specifically subjecthood, has often been argued to contribute to referent resolution (Crawley & Stevenson, 1990; Frederiksen,

---

and von Heusinger & Schumacher (2019) explicitly link accessibility / prominence and attention, Arnold & Lao (2015) found that visual capture cues on one of two depicted referents resulted in more attention towards that referent, but this effect did not always translate to the pronoun's interpretation, which led the authors to suggest that attention towards a referent does not necessarily equal a higher accessibility of that referent. However, the fact that the participants' pronoun interpretation was not influenced by the visual capture cues does not necessarily mean that the visual manipulation did not make the referent more prominent/accessible. It can also be interpreted as prominence not being determinant during pronoun interpretation; rather other factors like subjecthood and/or agentivity may be more important or reliable. Another question that needs to be addressed is whether the visual capture cue actually drew attention to the referent or whether attention was on the capture cue itself. Thus, the highlighting of one referent using visual cues may not have put the referent "in focus" (cf. Gundel et al., 1993). To that end, prosodic focus marking might be more suitable, as findings of Kristensen, Wang, Petersson & Hagoort (2013) suggest that prosodic focus marking engages general attention networks. Moreover, prosodic focus marking increases attention to the *concept* denoted by the expression relative to the linguistic context. Increased looks to a picture co-referential with the concept are due to this linguistic process. It is not at all clear that the same is achieved in a visual world the other way around, by highlighting (an aspect of) a picture co-referring to the concept, when it is not known whether and when a picture is in itself enough to activate the concept in question, let alone to focus it, in this paradigm (Magnuson, 2019)(also see discussion at the beginning of section 3.2).

1981; Fukumura & van Gompel, 2015; Gordon & Chan, 1995; Kaiser, 2011a; Song & Fisher, 2005). In support of order-of-mention, Gernsbacher & Hargreaves (1988) found the first mentioned referent to be at an advantage when compared to the second mention in a series of probe recognition tasks. Regardless of linguistic factors such as semantic or grammatical role, the response to the first mention was faster than to the second mention (see also e.g. Carreiras, Gernsbacher & Villa, 1995, for support of the first mention effect in Spanish). They argue that first mentioned referents are more available in a mental representation of the information being com- prehended. During the process of comprehension, this representation is built by laying a foundation based on the initially received information. Any subsequent information would be mapped onto this foundation (Gernsbacher, 1985; Gerns- bacher & Hargreaves, 1988; Gernsbacher et al., 1989) (the process of building a mental representation carries similarities with the organization of discourse struc- ture, see (e.g. Grosz & Sidner, 1986), and information structure, see e.g. (Krifka, 2008)). Thus, the first mention advantage is assumed to be connected to general cognitive processes that do not rely on linguistic factors like grammatical role.[2] However, Gernsbacher & Hargreaves (1988) did not investigate pronoun proces- sing. It is therefore not clear whether the advantage can be transferred to pronoun resolution.

In order to resolve this question, Fukumura & van Gompel (2015) proceeded to test the relative effects of order of mention and subjecthood in pronoun proces- sing compared to the processing of repeated names. They found that reference to the first-mentioned entity resulted in longer reading times than reference to the second-mentioned one, but only for repeated names (also see e.g. Almor, 1999; Gordon & Chan, 1995; Gordon et al., 1993; Shoji, Dubinsky & Almor, 2017, on the

---

[2]Of course, order-of-mention differences are often correlated with linguistic differences. For example, subject and object in an SVO sentence differ not only in order, but also in terms of semantic and grammatical role, as well as their positions in the syntactic structure. However, a major tenet of the first-mention account is that the effect itself is not linguistic. Therefore, crucially, it is assumed to hold also for referents differing only in order of mention, such as two subjects of consecutive SVO sentences or co-ordinated subjects and object in sentences such as *Alex and Lee bought books and magazines*. The claim that pronoun resolution is guided by a first-mention preference therefore means that the preference for subjects over objects is not due to their linguistic differences, i.e. it is not a linguistic preference.

repeated name penalty). For the pronouns, reading times were shorter when the referent was the subject than when it was the prepositional object. However, since in their stimulus set one of the referents was the topic (e.g. *Like Barry*[3] *Sally was in debt* is about Sally, making Sally the topic, see below for the definition of topics), this may have confounded the information structure and, consequently, the design. In other words, reference to Sally would (always) entail a pronoun, but in order to refer to Barry, a full NP would (always) be expected (and a pronoun would be penalized). While subjects and topics cannot be equated, there seems to be a strong correlation between them, where subjects will be interpreted as topics when no context is given (see Lambrecht, 1994, for further discussion). It is therefore likely that the role of the subject antecedent was unduly facilitated in Fukumura & van Gompel's study (conversely, see Bouma & Hopp's, (2006), criticism of Rambow's, (1993), argument for order of mention over subjecthood in German).

In order to disentangle grammatical role (subject vs object) and linear order (first- vs second-mention), previous research has often investigated pronoun resolution in languages with flexible word order, such as German and Finnish (Blything et al., 2021a; Bouma & Hopp, 2006; Järvikivi, van Gompel & Hyönä, 2017; Järvikivi et al., 2005; Kaiser, 2011a; Kaiser & Trueswell, 2008; Sauermann & Gagarina, 2017; Schumacher et al., 2016, 2017).

Research on Finnish subject pronoun resolution using SVO and OVS structures has shown either both subjecthood and first-mention effects (Järvikivi et al., 2005) or a subject-preference in the absence of a first-mention effect (Kaiser & Trueswell, 2008). However, as Järvikivi et al. (2005) presented their sentences in isolation, and thus did not control for possible changes in information structure for the two word orders (Kaiser & Trueswell, 2004), the observed first-mention preference could also be a preference for the given or discourse-old referent (topic). Further, as word order variation is used for other information structure functions in Finnish, in particular, OVS is also used to focus the subject antecedent, this could have highlighted

---

[3]It should also be noted that 'Barry', in contrast to the authors' claim, is not an object (oblique or otherwise). Instead, because the intended meaning of the sentence is not 'Sally is like Barry', 'like Barry' is either an adjunct PP or an ellipsis (cf. Sally is in debt like Barry [is in debt]). We thank Evangelia Daskalaki for discussion of possible syntactic analyses of this example.

the observed subject preference in (Kaiser & Trueswell, 2008).

As many different definitions have been put forward for concepts such as *focus*, *givenness*, or *topic* (see Krifka, 2008, for an overview of the notions of information structure), it is important to clarify them as they are used in the present study. For instance, common descriptions for focus are a sense of "newness", i.e. an update of the common ground (Lambrecht, 1994), the most important or prominent information (Dik et al., 1980), identifying the answer to a *wh*-question (Büring, 2003), or the indication of present alternatives which are relevant for the interpretation (Rooth, 1985, 1992). Givenness has been defined as information that has already been expressed and thus can be derived from the prior discourse (Schwarzschild, 1999). Finally, topic has been described as the "old" or "given" information of an utterance (Chafe, 1976), or in the sense of "aboutness", i.e. it refers to the entity about which information is given (Reinhart, 1981). However, in spite of these differences, identifying the information structure of a given sentence is usually not contentious. Specifically, while we espouse Reinhart's (1981) "aboutness" definition for topics, topics in the question-answer pairs used in the present study always coincide with "old information", as they are mentioned in the preceding discourse (but see e.g. Beaver, Clark, Flemming, Jaeger & Wolters, 2007; Büring, 2015; Féry & Ishihara, 2009, for more complicated cases involving old information as foci). Similarly, we constructed our items such that foci in the critical sentences always provide the answer to a *wh*-question (cf. Büring, 2003) and select one out of several possible alternatives (cf. Rooth, 1985, 1992), but are also new in the sense of not having been mentioned in the immediately preceding question.

Similar to Finnish, German allows for changes in word order, which are related to information structure (Frey, 2005, 2006; Kaiser & Trueswell, 2008; Vilkuna, 1995). But even with a relatively free word order, there is a robust preference for subject-first orders (e.g. Meng & Bader, 2000a; Schriefers et al., 1995). Processing OVS order seems to be more effortful than SVO order (Bahlmann, Rodriguez-Fornells, Rotte & Münte, 2007). However, factors such as appropriate context and parallel structure effects can decrease processing efforts of OVS order and level out its disadvantages with respect to SVO order (Kaiser & Trueswell, 2004; Weskott, 2003).

For example, the default information structure – without any special intonation – in German would be to place "given" or "old" information before "new" information (Fanselow, 2015; Neeleman & van de Koot, 2015). Thus, in SVO order the subject would be the given information, while in OVS order the object would be the given information per default. However, as has been previously observed (Frey, 2006), moving the object to the prefield[4] in German can lead to different information structures. The object can receive the role of the topic. Additionally, fronted objects as in (9) can express contrastive focus (Fanselow, 2015; Frey, 2005, 2006) or appear in sentences with broad focus (answering *what happened?*) (Büring, 1997; Jacobs, 1991).

(9) Den   Arzt   hat der   Feuerwehrmann gerettet.
     the.ACC doctor has the.NOM firefighter   saved
     'The firefighter has saved the doctor.'

Thus, the non-canonical word order does not automatically change the object constituent's marking from focus to topic (Sauermann & Gagarina, 2017). Similarly, in English, Gundel (1974), (as cited in Reinhart, 1981, 63), observed that fronted constituents (e.g. via PP preposing) can be either topics or foci. When marked with special intonation, they can be interpreted as focused, however when no special intonation is present, fronted constituents are to be interpreted as topics. As a result, when sentences are presented out of context and / or prosody to guide the information structure assignment, more than one information structure can be assigned even for marked word orders. This highlights the importance of controlling information structure in studies of phenomena such as pronoun resolution, where it has been argued to have potential effects (see discussion below).

Sauermann & Gagarina (2017) manipulated word order and grammatical role to investigate German pronoun processing in an eye-tracking study. The gaze data showed a subject preference for the subject pronoun for both SVO and OVS word orders. The authors did not discuss any possible effects word order changes could have on information structure nor did they provide information on their items'

---

[4]In German the position preceding the finite verb in verb-second clauses is known as the *Vorfeld* or prefield (e.g. Frey, 2005, 2006).

intonation. They used a context sentence before the word order manipulation, but it may not have been enough to keep order of mention and information structure constant across the two different word orders. This means their results could be interpreted differently. We could assume that information structure coincided with grammatical role, i.e. the subject coincided with topic and the object with focus. Thus, the subject effect would simultaneously be a topic effect.

While information structure effects are often not considered in studies manipulating word order to disentangle subjecthood and first mention effects on pronoun resolution, another line of research has directly investigated information structure effects on pronoun resolution.

Most of these studies have manipulated syntactic means of focusing by using it-clefts, with different results. Some studies suggest that participants are more likely to attach personal pronouns to the subject antecedent when it is clefted than without clefting (Colonna et al., 2015; Foraker & McElree, 2007, e.g. *It was the actor who called the cook. He...* vs *The actor called the cook. He...]*, cf.¿[). However, others report either no difference in resolution preferences between clefted and non-clefted (SVO) antecedents, especially subjects (Colonna et al., 2014; Järvikivi et al., 2014; no difference between non-focused, topicalized and cleft subjects: Cowles et al., 2007; only a marginal effect of clefting: Kaiser, 2011) or an "anti-focus effect", i.e. fewer choices of clefted than non-clefted subject referents (Colonna et al., 2012, 2015; De la Fuente & Hemforth, 2013; Patterson et al., 2017).

However, most of these studies lack a felicitous context that would license the use of focus marking (but see De la Fuente & Hemforth, 2013; Kaiser, 2011a).

Similarly, the effect of topicality on pronoun resolution has mostly been investigated with out-of-the-blue sentences using syntactic topic marking. For instance, Colonna et al. (2012) tested left-dislocations (e.g. *As for Peter...*) for French and German, and De la Fuente & Hemforth (2013) for Spanish. They found that topicalization enhances an already existing referent preference (subject referent in German, object referent in French and Spanish).

Further, while these studies largely concentrate on syntactic marking, information structure influences different linguistic domains. In terms of prosody and

intonation, it affects the location and types of pitch accents in many languages (e.g. Féry, 2017; Jun, 2005, 2014; Kügler & Calhoun, 2020). In German, focus is as-sociated with a falling accent (H*(+L)) while post-focal given information is deac-cented (e.g. Baumann, 2006; Büring, 1997; Féry, 1993). Topics, on the other hand, are connected to rising accents (L*+H), especially when contrastive (e.g. Braun, 2006; Büring, 1997; Féry, 1993; Repp & Drenhaus, 2015). In various studies on the effects of intonation in online processing during eyetracking (Chen, Den Os & De Ruiter, 2007; Ito, Bibyk, Wagner & Speer, 2014; Ito & Speer, 2008; Sedivy, Tanenhaus, Chambers & Carlson, 1999), contrastive focus marking has been found to trigger anticipatory eye-movements to alternatives, e.g. anticipation that the in-coming noun will be of the same kind as the preceding one like in *blue ball* fol-lowed by *GREY ball* (Ito et al., 2014; Ito & Speer, 2008), which in turn can aid in target search. Similarly, initial higher proportions of looks towards the competitor have been found to decrease earlier for H*L (focus) conditions than for L*H (non-focus) conditions (Chen et al., 2007; Sedivy et al., 1999). In spite of these known effects, the role of prosody is generally not addressed in studies on the effect of information structure on pronoun resolution, whether it be the intonation readers imagine for written stimuli or the prosody accompanying syntactic manipulations in aurally presented stimuli.

Finally, it is not clear to what extent thematic, and not syntactic, roles determine pronoun resolution. Research investigating the role of thematic role in pronoun resolution found that with implicit causality verbs, thematic role outranks grammatical role with respect to the preferred referent (e.g. Rohde & Kehler, 2014; Stevenson, Crawley & Kleinman, 1994). Järvikivi et al. (2017) come to a similar conclusion with respect to implicit causality verbs in Finnish. Leaving implicit causality aside, Schumacher et al. (2016, 2017) used dative-experiencer verbs to tease apart (proto) agent from subject role and found the personal pronoun to be more often resolved towards the (proto) agent than the subject. We will return to this in the General Discussion.

In sum, the studies summarized above suggest that both focus and topicality may play a role in referent preference for ambiguous pronouns. However,

results vary according to various factors such as the language investigated, but also depending on whether or not context was used to establish a felicitous use of information structure manipulation. Most studies have looked only at how syntactic focus and topic marking affect pronoun resolution, but have not considered prosodic effects. Similarly, only few studies have taken information structural effects of word order changes into account when trying to disentangle subject and first mention effects on pronoun resolution. Moreover, the subject role is often confounded with first-mention, topicality, and agentivity, and indeed most researchers assume that multiple factors influence pronoun resolution (e.g. Arnold, 2010; Gordon et al., 1993; Järvikivi et al., 2005; Kaiser & Trueswell, 2008).

The current study tests the role of order-of-mention, grammatical role, and information structure on pronoun resolution, addressing the above-mentioned research gaps regarding prosody and information structure using prosodic focus marking and word order manipulation. In addition, using of eye-tracking, we analyze the role of prominence/accessibility during referent and pronoun processing.

3.2 Current Study

In Experiment 1, we used the *Visual World Paradigm* to provide a sensitive tool to assess how participants' attention towards the depicted referents would change in real-time with changes in prosody and word order, as well as assessing their real-time referent preference for the ambiguous pronoun. The use of this paradigm is grounded in literature showing that listeners look toward an element depicted on the screen upon hearing it mentioned in the input (Altmann & Kamide, 1999; Järvikivi et al., 2005). In the present study, we evaluated participants' eye movements not only during the time segment when they heard the pronoun, as is commonly done, but additionally also during the immediately preceding context where the potential pronoun referents were mentioned. Thus, using the *Visual World Paradigm* will not only give us information on where participants looked when hearing the pronoun, but also where they looked when hearing the referents. Knowing where participants looked when hearing the referents will shed light on whether

accessibility / prominence of a referent is indeed the critical factor for pronoun resolution as we can use this as a measure of attention. This is useful, since referents that receive more attention are said to be more accessible for pronouns (Gundel et al., 1993), resulting in a higher possibility for a referential link to be established. In the Experiment 1, we manipulated word order to tease apart **subjecthood** from **first-mention** effects in German (see Table 3.1 for example). However, object fronting can cause differences in information structure; as previously noted, it can either signal that the fronted object is a topic or that it is a contrastive focus in German OVS sentences. In order to resolve this ambiguity and to exclude any additional processing effects that could stem from OVS order, we embedded the manipulation into appropriate contexts. Further, we manipulated **information structure** by applying prosodic focus and given topic marking on subject and object referents. We asked whether the previously found robust subject preference can be modulated by prosodically marking the possible referents. In Experiment 2, we looked at how the personal pronoun *er* and the demonstrative pronoun *der* differed regarding the effects of grammatical role and prosodic focus marking for referent selection.

In our experiments, the marking of focus and topic was twofold, using matching contexts and pitch accents to mark one referent with focus, while the other referent was established as the (given) topic via context as well as remaining unaccented. Although coupling context-induced focus (via *wh*-questions) with prosodic focus marking has not been found to significantly increase processing advantage for the focused entity when compared to only one type of focus marking being present (e.g. Akker & Cutler, 2003; Cutler & Fodor, 1979), marking focus additionally with a pitch accent can still be useful when processing utterances as it can guide listener's attention to the "prominent" or "important" constituents (Féry, 2017, p. 142). Moreover, as mentioned above, even marked OVS orders can have more than one possible information structure in German, which prosodic focus marking allows to disambiguate. It also constitutes the most natural realization of focused material.

As mentioned above, the analysis of the gaze data (Experiment 1) is split into

two time segments. The first segment included the prosody and word order manipulation of the two intended referents (subject and object), the second segment contained the personal pronoun *er*. Generally, in eye-tracking research on pronoun resolution, only the segment that includes the pronoun is taken into account. So, we do not know how the pronoun's possible referents are processed. Including the segment preceding the pronoun has a twofold aim. First, since this is the segment that contains the prosodic manipulations of the referents (focus and given topic), we want to make sure that these manipulations actually affected participants' gaze behavior. Second, since it has been claimed that prominent referents receive more attention, this prominence should also be reflected in participants' gaze data. Based on previous research that focus marking engaged general attention (e.g. Kristensen et al., 2013; Osaka, Komori, Morishita & Osaka, 2012), we assume that this translates to the focus marked referent receiving more looks than the non-focused referent.

For the second segment containing the ambiguous pronoun, if **subjecthood** drives pronoun resolution, we predict more looks to the subject referent regardless of information structure and order of mention. However, if **information structure** drives pronoun resolution, we expect more looks to the focus-marked referent regardless of grammatical role, a pattern similar to the first segment. This result would also support accessibility / prominence-based accounts of pronoun resolution, especially if focused referents receive increased looks during the first segment, in line with focus marking attracting participants' attention to the marked referents. If **first-mention** is the crucial factor, we predict the first-mentioned referent to receive more looks than the second-mention referent regardless of grammatical role and information structure. As prior research also found these factors to interact with each other, we could expect that when first-mention, subject role and focus marking align then that referent will receive more looks compared to a referent where these factors to not align.

For the final pronoun interpretation, since previous findings suggest a strong **subject** referent preference for the personal pronoun *er* 'he', we expect this preference to hold here as well, but it may be modulated by information structure or

first-mention bias. If **information structure** plays a role, we predict to see more subject referent choices in the subject focus condition than in the object focus condition. In the object focus condition, the effect of subjecthood against the focus effect may make referent preference less clear, which could possibly result in referent choice being at chance level. Again, if participants paid increased attention to focused referents during Segment 1, any effects of information structure on the final interpretation may be taken as support for the idea that pronouns prefer the most accessible, salient or prominent referent, and that attention can be used to substantiate the notion of accessibility/salience/prominence. If **order of mention** affects the subject preference, we expect more subject choices in SVO order than in OVS order, and more object choices in OVS than in SVO order. As several studies found that referent preferences can be clearer when several effects combine (e.g. Blything et al., 2021a; Kaiser & Trueswell, 2008), it is also possible that the subjecthood effect is modulated by information structure and/or word order.

Predictions for the demonstrative pronoun *der* will be discussed in section 3.5.2.

3.3 Data Availability

The stimuli, data, scripts, and additional plots are available at:

```
https://osf.io/athvb/?view_only=9022d5258e0a48328eef9faaf458f205
```

3.4 Methods

3.4.1 Participants

Sixty German-native speakers from the University of Konstanz and the University of Oldenburg completed the experiment (44 female, 15 male, 1 non-binary; age range: 20-33, mean: 24.5, sd: 2.78). All participants reported normal hearing and normal or corrected-to-normal vision. They were given monetary compensation (8-10 €) for their participation.

The study was approved by the Research Ethics Board 2 of the University of Alberta (study ID Pro00076383).

3.4.2 Materials
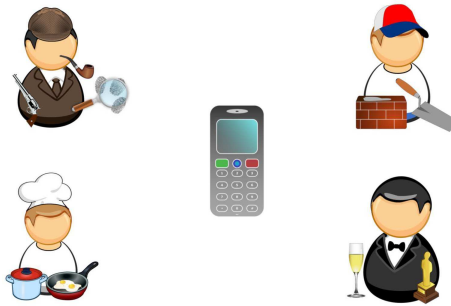
Figure 3.1: Sample of visual stimuli during eye-tracking showing the four referents from the dialogue as well as the instrument/object that is mentioned before the pronoun.

The experimental items consisted of 40 dialogues which varied across four lists in word order (SVO, OVS) and prosodic focus marking (subject focus, object focus), see example dialogue in Table 3.1. The use of dialogues helps to create an appropriate context to control information structure and decrease processing efforts connected to OVS order as mentioned earlier. The first sentence of the dialogue introduced four referents, who were referred to by their profession. For the referents in the critical sentence (prosody and word order manipulation), the number of syllables varied: 3 monosyllabic, 13 disyllabic, 14 trisyllabic, 9 quadrisyllabic, and 1 of six syllables. For the majority, lexical stress was on the first syllable, the penultimate syllable was stressed in 8, and the ultimate syllable was stressed in 5 referents. We accounted for item variation in our statistical models.

As mentioned in section 3.2, focus and topic marking was twofold. Context-induced focus marking was coupled with prosodic focus marking.

In order to render the use of the prosodic marking felicitous, we added a *wh-*question that preceded the critical word order and prosodic marking manipulation, which established the appropriate information structure for the critical sentence. Preceding subject questions were used to establish the subject referent as focused (and the object as a given topic), while object questions established the object referent as the focus (and the subject as a given topic). In the critical sentence, prosody marked one constituent as focused and one as a given topic. The condition name indicates the role of the focused constituent, for instance, in the

subject focus condition the subject was in focus while the object was marked as the given topic. This manipulation was crossed with manipulation of word order in the critical sentence in a 2 × 2 design.

A sentence containing the unaccented and ambiguous subject pronoun *er* followed the sentence with the word order and prosodic marking manipulation after a pause of 1000 ms (pauses of the same length appeared after all units marked by paragraph breaks in Table 3.1).

The dialogues were recorded using Shure SM10A headset microphones in a sound attenuated booth by two native speakers of German, one female and one male. The female speaker (the last author, a prosody researcher) recorded all the introduction and critical sentences (A-turns in Table 3.1) for the experimental items while for the filler items it was the male speaker. Figure 3.2 visualizes representative pitch contours for all experimental conditions. The speaker uniformly produced a single falling accent on the focused constituent, while the rest of the sentence remained unaccented (represented as H* L-% in GToBI notation Grice, Baumann & Benzmüller, 2005; Grice, Baumann, Ritter & Röhr, 2017). Thus, in the conditions where the first constituent was focused (subject focus in SVO, cf. Figure 3.2a, and object focus in OVS, cf. Figure 3.2b), this first constituent carried a pitch peak followed by a fall and a low plateau stretching until the end of the sentence. In the other two conditions (object focus in SVO, cf. Figure 3.2c, and subject focus in OVS, cf. Figure 3.2d), the pitch on the first constituent was low followed by a rise to a peak on the second (focused) argument of the verb and a final pitch fall (additional information on the prosody manipulation, such as plots for average contours, duration, and intensity are available on `https://osf.io/athvb/?view_only=9022d5258e0a48328eef9faaf458f205`).

The experimental items were distributed across four lists in a Latin square design, each containing 10 items per experimental condition. Additionally, we constructed 40 filler items which were the same across all lists. These fillers also contained four possible referents, but unlike the experimental items we did not include any ambiguous pronouns.

(a) Subject focus in SVO.

(b) Object focus in OVS.

(c) Object focus in SVO.

(d) Subject focus in OVS.
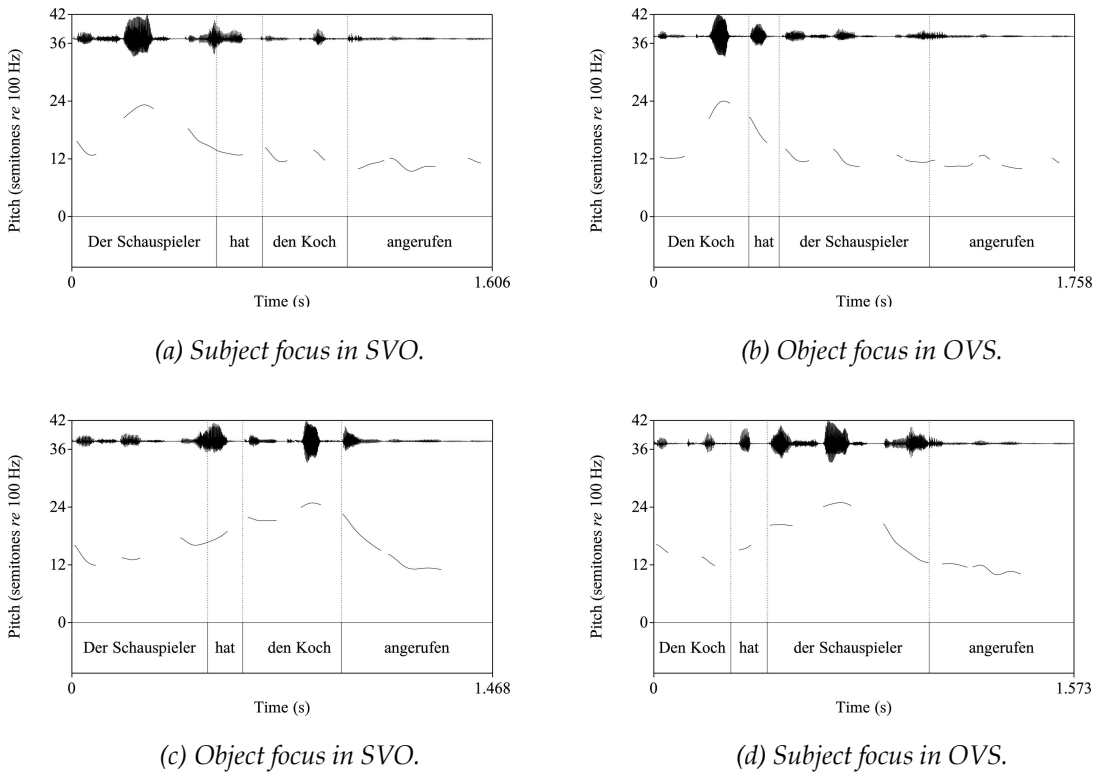
Figure 3.2: Prosodic contours across different conditions.

3.4.3 Procedure

The experiment took place in a quiet room. The participants were given a brief oral explanation of the tasks they were about to complete. A more detailed instruction for the eye tracking task would be seen on the screen after the calibration. We used an SR Research Eyelink Portable Duo eye tracker and recorded with a sampling frequency of 500 Hz in remote head free-to-move mode. For all participants, only the right eye was tracked. We opted for a 9-point calibration. After the calibration, a screen with instructions appeared, asking participants to carefully listen to the dialogues. While listening to the dialogues, participants were presented pictures of the four mentioned referents and one instrument (see Figure 3.1) which was mentioned in the dialogue between the critical sentence and the ambiguous pronoun (e.g. a mobile phone for the dialogue in Table 3.1). Following each dialogue, they saw a question on the screen probing to which of the two target referents, subject or object, the pronoun referred (see last row in Table 3.1). We also included

the other two referents as possible responses to ensure that participants paid attention during the experiment. Participants gave their answer using one of four keys on a Cedrus response pad RB-844 that mirrored the position of the pictures on the screen. The positions of the referents on the screen were randomized for each trial. The stimuli were presented visually on an external 21″ screen and auditorily with a Bose SoundLink mini speaker. Halfway through the experiment participants were given a break, which was then followed by a second calibration.

3.5 Experiment 1: Eye-tracking and referent selection

3.5.1 Offline data: Referent selection

We first present the offline response data reflecting participants' final interpretation of the ambiguous pronoun *er*.

Analysis

We performed generalized linear mixed-effects regression modeling (GLMER) using the lme4 package (version 1.1-23, Bates et al., 2015) in the software R (version 3.6.3, R Core Team, 2019) to analyze the offline responses. The model included a binomial dependent variable coding whether the participant chose the subject or the object as the referent of the pronoun (*SOPref*). Responses choosing one of the distractor referents were discarded from modelling (N=58, 2.4% of responses; leaving 2383 data points for analysis). We included an interaction between *Focus* and *WordOrder*, a by-participant random slope for *Focus*, a by-participant random slope for *WordOrder*, as well as a random intercept for *Item*. Aside from the intercept, neither *WordOrder* nor *Focus*, nor the interaction between these two effects were significant. We compared this model to a simpler model which excluded the interaction, but was identical otherwise. An improvement was assessed by using the likelihood ratio test (Matuschek, Kliegl, Vasishth, Baayen & Bates, 2017). As the interaction did not improve the model's fit, it was consequently dropped in the final model.

Results

Figure 3.3 shows the referent choices for the pronoun *er* by word order (SVO, OVS) and focus (subject focus, object focus). As the Figure shows, there was an overall preference for the subject as the referent of the ambiguous pronoun. The best model (see Table 3.2) revealed a subject preference (positive intercept). In addition, we compared the overall subject preference as well as the subject preference per word order and prosody condition to chance level using one-sample Wilcoxon signed rank test and found the difference to be significant (for all p<0.05). The effect of focus was marginal in the best model, but it was not significant due to participant and item variance. The effect of word order was not significant.
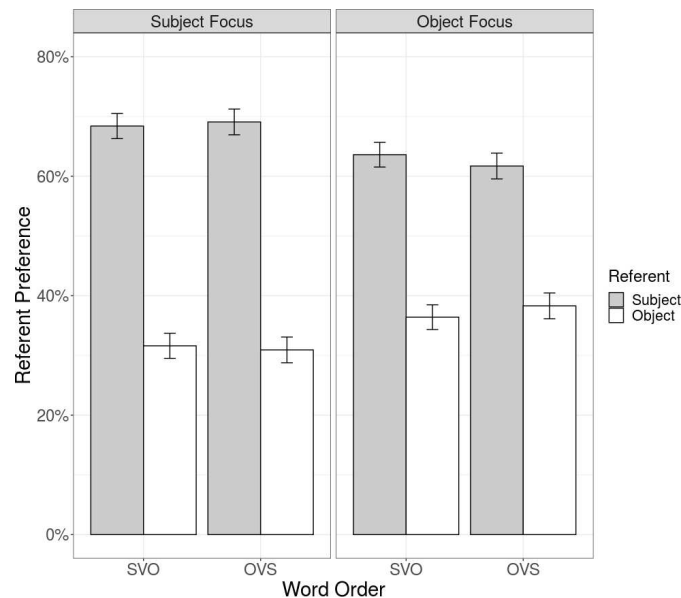
Figure 3.3: Referent choice for er, in percent, by word order and prosody in the offline response with error bars for standard error.

3.5.2 Eye-Tracking Data

Analysis

The eye gaze data was split into two separate analysis windows. The first one contained the critical sentence with the word order and prosodic manipulation (below: "Segment 1"), the second one included the sentence containing the ambiguous pronoun *er* ("Segment 2"). The gaze data was preprocessed in R using the

VWPre package (version 1.2.2, Porretta, Kyrö läinen, van Rij & Järvikivi, 2016). The data points were first binned into 20 ms bins and, in a second step, transformed into empirical logits (Barr, 2008; Porretta et al., 2018).

For the statistical analysis, we used generalized additive mixed-effects modeling (GAMM) in R using the package mgcv (version 1.8-38, Wood, 2017). GAMM is especially useful for analysing time series data (Thul, Conklin & Barr, 2021), like visual world eye tracking data, because it does not entail averaging over time windows or assume that relationships between continuous variables are linear (van Rij et al., 2016). GAMM allows for the inclusion of non-linear relationships using smooths (Wieling, 2018). With GAMMs, visualization is important in the process of model evaluation and for the interpretation of the model's estimated effects (Porretta et al., 2018). We therefore used the the tools provided by the itsadug package (version 2.3, van Rij et al., 2020b). The dependent variable, *SODifference*, was created by subtracting the eLogit transformed values of looks to object from looks to subject referent. In GAMMs, combined factors are used to model interactions that involve more than one categorical predictor and one or more continuous predictors (e.g. van Rij et al., 2016). Thus, in order to determine whether there were differences for word order for each type of focus condition over time, we combined *WordOrder* with *Focus* into a predictor *F WO* with four levels. All models included *SODifference* as the response variable, the combined factor *Focus WO* with four levels, which was included as a **parametric term**, as well as smooth terms for the interaction between the combined predictor *Focus WO* with *Time*. Further, the model included random intercepts for *Event* (a unique combination of *Participant* and *Trial*), and random smooths for *Participant* and *Item* to account for individual variation between participants. For Segment 2, we also included an AR1 (*autoregessive*) parameter in the model to account for autocorrelated errors in the data (e.g. Baayen et al., 2018; Porretta et al., 2018). The AR1 model specified the autocorrelation coefficient *rho* and the starting point for each time series (Baayen et al., 2018). This was not needed for Segment 1 since the residuals showed hardly any autocorrelation to begin with.

Results

Gaze Data: Segment 1

Figure 3.4 shows the grand average proportion of looks over time for both word order and focus conditions for Segment 1, which consisted of the critical sentence with the prosody and word order manipulation. In this segment, the focused referent received more looks over time in all four conditions. For the conditions where the subject referent was in focus, there was an initial increase in looks towards the subject which later decreased with time, while for the object focus conditions this pattern was mirrored in the looks to the object referent. When the focused referent was first-mention, the peak in looks to the focused referents was earlier (for SVO Subject Focus 1260 ms and for OVS Object Focus 1220 ms) than when it was second-mention (for OVS Subject Focus 1900 ms and for SVO Object Fo- cus 1620 ms). Interestingly, the difference in time for the peaks is relatively small, which could suggest a prosody effect that resulted in a relatively fast shift of atten- tion towards the focused referent even when the focused referent was the second- mention.

Table 3.3 represents the summary of the inferential statistics for the optimum-fit model for Segment 1. For the parametric coefficients (here time course is not taken into account), the *p*-value reveals whether a combination of word order and focus is significantly different from the reference level, subject focus + SVO. Positive estimate values indicate more looks to the subject referent overall, while negative values signal more looks to the object referent. Both object focus conditions were significantly different from the intercept, showing a decreased preference for the subject referent. Subject focus + OVS did not differ significantly from subject focus + SVO, meaning that changes in word order did not cause changes in looks towards the referents.

For the smooth terms, the *p*-value indicates whether or not a smooth is signif- icantly different from 0. The *edf* value represents the number of effective degrees of freedom which estimates the number of parameters needed for the smooth, and reflects the degree to which the pattern is (non)linear. A value of 1 corresponds to
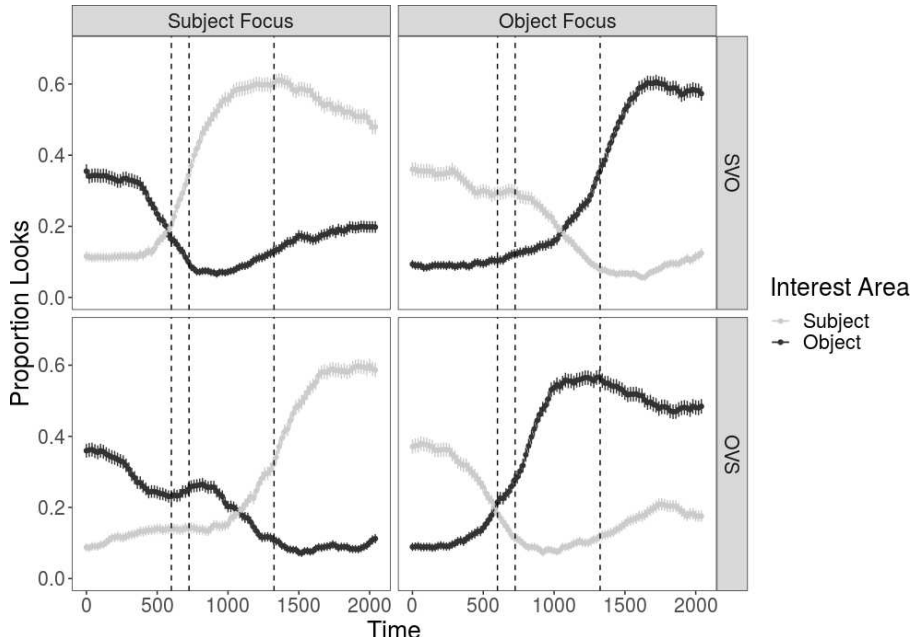
Figure 3.4: Grand average proportion of looks to subject and object by condition in the first segment. The x-axis shows time from the beginning to the average end of Segment 1. 0 ms marks the onset of the first referent, first dashed line marks marks average onset of auxiliary, second dashed line marks average onset of the second referent, third dashed line marks average onset of participle.

a linear pattern while a value greater than 1 suggests a more complex non-linear pattern. For this segment, all smooth terms were significantly different from 0 and non-linear (edf > 1).

To visualize the model's predictions and to determine whether smooths of each experimental condition significantly differed from each other, difference plots are needed. Figure 3.5 presents these difference plots for Segment 1, showing the differences between six pairs of focus and word order condition combinations over time. On the y-axis are the estimated differences between the looks to the subject and looks to the object. A positive value indicates more looks towards the subject, while a negative value indicates more looks towards the object for the comparison of the first and second listed conditions in the caption and title of each plot. On the x-axis, 0 ms is the onset of the critical sentence. Significant differences between compared conditions are marked in red.

For Figure 3.5a, the comparison we observe is between subject focus and object focus in SVO order. The first significant time window (0-686 ms) shows more

looks towards the object referent for the subject focus condition when compared to the object focus condition, though the difference is rapidly decreasing. This initial difference is likely a spillover from the previous question in the dialogue: For the subject focus condition, the last referent participants heard mentioned was the object, while in object focus condition the last referent participants heard was the subject. From 768 ms onward, this changed to more looks towards the subject for subject focus compared to the object focus condition, suggesting that focus guides listeners' attention to the focused referent. The same pattern also appeared for OVS order: Here, too, the object referent initially received more looks in subject focus than in the object focus condition (0-646 ms), but from 767 ms the subject received more looks in subject focus than in object focus condition (Figure 3.5b).

The next two plots compare subject and object focus on the first mention referent (Figure 3.5c) and on the second mention referent (Figure 3.5d). As illustrated in Figure 3.5c, focus on the first-mentioned subject in SVO showed initially more looks to the object than focus on the first-mentioned object in OVS for a short time window (0-566 ms). This preference then (646-4000 ms) switched to more looks towards the subject in subject focus SVO compared to object focus OVS. Figure 3.5d showed more looks towards the object in subject focus OVS than in object focus SVO word order first (0-1010 ms), then (1090-4000 ms) this changed to more looks towards the subject for subject focus OVS than for object focus SVO. Both comparisons robustly showed more looks to the subject in subject focus and fewer looks to the subject in object focus across the majority of the time window. As is to be expected, this significant subject preference in the subject focus condition started earlier when comparing first-mentioned focused referents (646 ms in Figure 3.5c) than when comparing second-mentioned focused referents (1090 ms in Figure 3.5d).

The last two plots show the difference between subject focus in SVO and OVS order (Figure 3.5e) and between object focus in SVO and OVS order (Figure 3.5f). The conditions in which the subject was focused (Figure 3.5e) showed more looks to the subject referent in SVO than in OVS order first (525-1495 ms), followed by more looks towards the object in SVO than in OVS order (1616-3879 ms). Thus,

when focus was kept constant to the subject position, there were more looks to the subject referent at an earlier time in SVO, since the first mention was the subject, compared to OVS, where the subject was second mentioned. Comparing the conditions in which the object referent was focused (Figure 3.5f), from 485 ms to 1414 ms, there were more looks towards the subject in object focus SVO than in object focus OVS. From 1535 ms to 3111 ms, looks towards the object increased significantly. Again, when focus was kept constant (to the object position), there were initially more looks to the subject referent in SVO compared to OVS, since the first mention was the subject.

In sum, prosodically focus-marked referents received more looks than referents that were marked prosodically as given topics. Neither grammatical role nor word order as such affected proportion of looks as can be also seen in Figure 3.4. As outlined in section 3.2, we assumed that focus marked referents would receive more looks than non-focus marked referents, which was based on findings that focus marking engages general attention (Kristensen et al., 2013; Osaka et al., 2012). Therefore, we take the focus marked referent to receive more attention than the non-focus marked referent which – under the assumption that (focus driven) attention correlates with prominence / accessibility, as discussed in section 3.1 – means that the focus marked referent was more prominent / accessible than the non-focus-referent.

Gaze Data: Segment 2

Figure 3.6 shows the grand average proportion of looks over time for both word order and focus conditions for the second segment, which started with the ambiguous pronoun. In Segment 2, looks to the subject referent increased from the onset of the pronoun in the two subject focus conditions, whereas for the object focus conditions looks were initially higher for the object referent but then switched around 1500 ms after the onset of the pronoun to more looks towards the subject.

Table 3.4 represents the summary of the inferential statistics for the optimum-fit model. For the parametric coefficients, similar to the first segment, object focus + SVO and object focus + OVS were significantly different from the reference level

subject focus + SVO. Both showed fewer looks to the subject referent than the intercept overall. Subject focus + OVS again was not significantly different from subject focus + SVO. For the smooth terms, both subject focus conditions were significantly different from 0, as was the object focus + OVS condition.

Table 3.1: Example dialogue in SVO and OVS word order conditions. Sentence with critical manipulation of word order and information structure in italics, unaccented ambiguous pronoun in bold. Contexts were identical for all conditions except where indicated with slashes and condition names in brackets. Note that word order was manipulated within the critical sentence, while all other parts of the dialogue did not differ between word order conditions. Information structure was manipulated prosodically in the critical sentence (not indicated here), as well as in the preceding context.

German	English translation
A: Ich habe gerade Ärger in meiner Strickgruppe, in der auch der Koch, der Schauspieler, der Maurer und der Detektiv sind. Wir haben einen Termin verschoben und ziemlich viel rumtelefoniert. Als letztes hat jemand den Koch angerufen. (Subject focus) / Als letztes hat der Schauspieler jemanden angerufen. (Object focus)	A: I have some problems in my knitting group which also includes the cook, the actor, the bricklayer, and detective. We postponed an appointment and called back and forth. Lastly, someone called the cook. (Subject focus) / Lastly, the actor called someone. (Object focus)
B: Und wer hat den Koch angerufen? (Subject focus) / Und wen hat der Schauspieler angerufen? (Object focus)	B: And who called the cook? (Subject focus) / And whom did the actor call? (Object focus)
A: *Der Schaupieler hat den Koch angerufen* (SVO), / *Den Koch hat der Schauspieler angerufen* (OVS), und zwar mit einem Handy. **Er** war zu diesem Zeitpunkt schon ziemlich m ü de.	A: The.NOM actor has the.ACC cook called (SVO), / The.ACC cook has the.NOM actor called (OVS), *The actor called the cook,* namely with a mobile phone. **He** was already pretty tired at this point.
B: Das ist aber schade.	B: That is too bad.
Q: Wer war schon ziemlich müde?	Q: Who was already pretty tired?

86

Table 3.2: Fixed effects for best fitting generalized linear mixed-effects model of referent choice.

Formula: SOPref ~ Focus + WordOrder + (1 + Focus | Participant) + (1 + WordOrder | Participant) + (1 + Focus | Item)+ (1 + WordOrder | Item), control=glmerControl (optimizer ="bobyqa")

	Estimate	Std. Error	z value	Pr(| > z|)
(Intercept)	1.1852	0.2711	4.371	1.23e-05
Focus Object Focus	-0.2654	0.1435	-1.849	0.0644
WordOrder OVS	-0.2119	0.1490	-1.422	0.1550

Table 3.3: Model summary for gaze data for Segment 1.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	1.2929	0.1077	12.0089	< 0.0001
Focus_WOSubjectFoc_OVS	-0.0104	0.1396	-0.0746	0.9405
Focus_WOObjectFoc_SVO	-2.6726	0.1397	-19.1376	< 0.0001
Focus_WOObjectFoc_OVS	-2.7594	0.1394	-19.7969	< 0.0001
B. smooth terms	edf	Ref.df	F-value	p-value
s(Time):Focus_WOSubjectFoc_SVO	8.8746	8.9282	69.8794	< 0.0001
s(Time):Focus_WOSubjectFoc_OVS	8.7197	8.8392	45.6508	< 0.0001
s(Time):Focus_WOObjectFoc_SVO	8.6847	8.8385	57.8113	< 0.0001
s(Time):Focus_WOObjectFoc_OVS	8.8082	8.8856	57.7646	< 0.0001
s(Event)	1251.4058	1295.0000	48.0817	< 0.0001
s(Time,Participant)	245.4684	296.0000	37.5884	< 0.0001
s(Time,Item)	284.5594	359.0000	71.6178	0.0003

Table 3.4: Model summary for gaze data for Segment 2.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.5299	0.1454	3.6435	0.0003
Focus_WOSubjectFoc_OVS	0.0772	0.1395	0.5532	0.5801
Focus_WOObjectFoc_SVO	-0.6860	0.1392	-4.9267	< 0.0001
Focus_WOObjectFoc_OVS	-0.4868	0.1397	-3.4847	0.0005
B. smooth terms	edf	Ref.df	F-value	p-value
s(Time):Focus_WOSubjectFoc_SVO	7.8610	8.5440	17.7554	< 0.0001
s(Time):Focus_WOSubjectFoc_OVS	7.2244	8.1074	35.0977	< 0.0001
s(Time):Focus_WOObjectFoc_SVO	1.0028	1.0034	0.0286	0.8693
s(Time):Focus_WOObjectFoc_OVS	8.2955	8.7851	6.5998	< 0.0001
s(Event)	2189.6205	2362.0000	57.2293	< 0.0001
s(Time,Participant)	402.0340	548.0000	40.8431	< 0.0001
s(Time,Item)	258.0315	359.0000	774.9509	< 0.0001
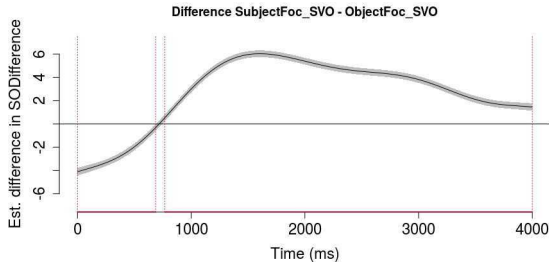
(a) Difference in looks between subject focus and object focus in SVO
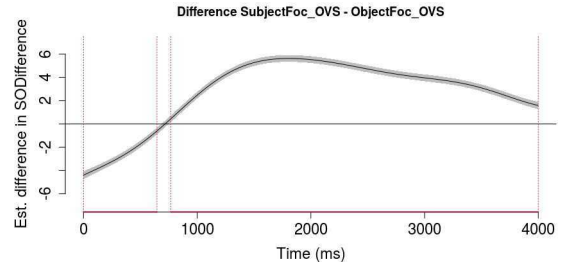
(b) Difference in looks between subject focus and object focus in OVS
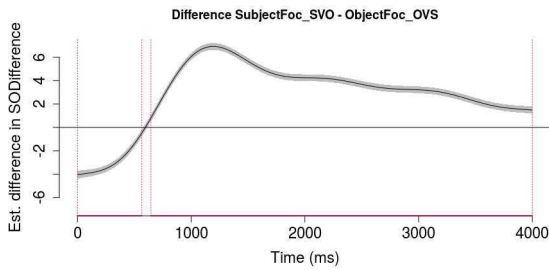
(c) Difference in looks between between subject focus SVO and object focus OVS
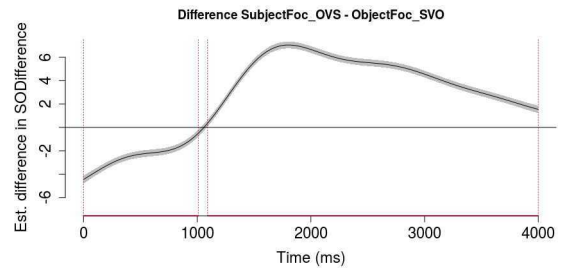
(d) Difference in looks between between subject focus OVS and object focus SVO
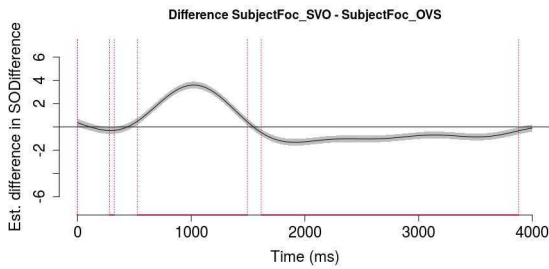
(e) Difference in looks between between subject focus in SVO and OVS
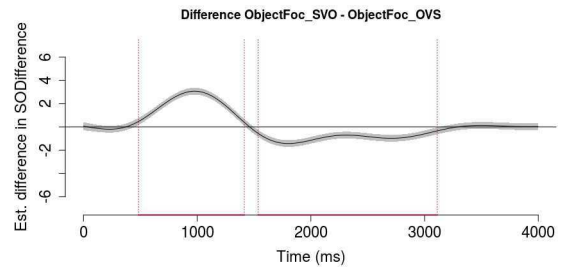
(f) Difference in looks between between object focus in SVO and OVS

Figure 3.5: Difference plots across all four conditions for the segment with the prosodic and word order manipulation. Shaded areas represent 95% confidence interval. Black line indicates zero effect.
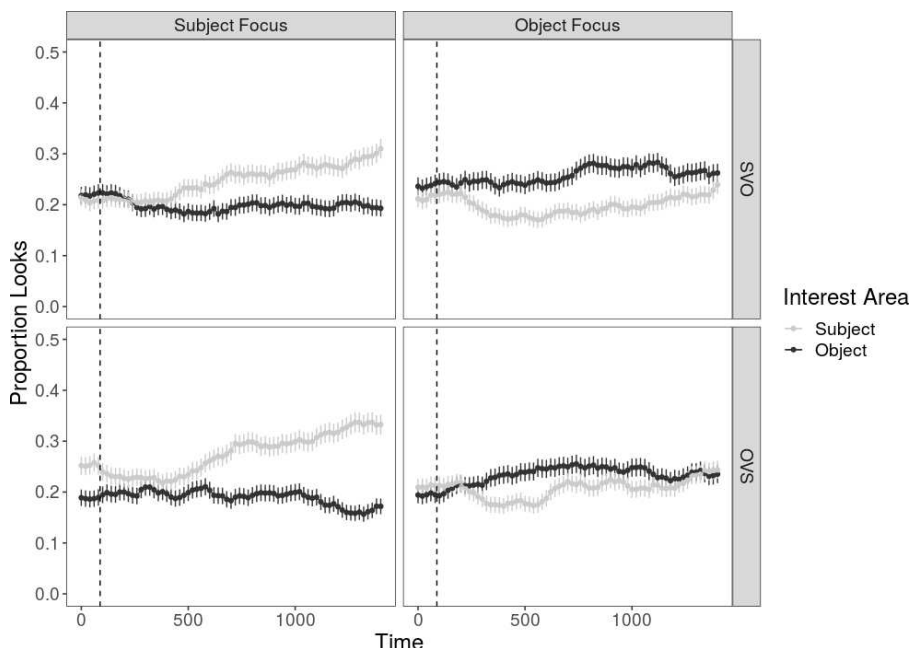
Figure 3.6: Grand average proportion of looks to subject and object by condition in the second segment. The x-axis shows time from the beginning to the average end of Segment 2. 0 ms is the onset of the pronoun, dashed line is average onset of the verb.

Difference plots showed significant differences in the following combinations: First, when keeping word order constant to SVO (Figure 3.7a), there were more looks to the subject referent for subject focus than for object focus starting shortly after the onset of the pronoun (182 ms) until the end of the segment. When keeping word order constant to OVS (Figure 3.7b), from 273 ms until 2273 ms the subject received more looks in the subject focus than in the object focus condition. After 2600 ms, this shifted to more looks towards the the object referent in subject focus than in object focus. Similar to the previous Segment 1, this segment also showed an overall preference for the focused referent.

When comparing subject and object focus on the first mention referent (Figure 3.7c) and on the second mention referent (Figure 3.7d) the subject referent received more looks in the subject focus condition compared to the object focus condition (303-2333 ms and 2879-3000 ms for first mention, and 0-2485 ms for second mention).

Comparing the two conditions in which the subject referent was focused (Figure 3.7e), the first significant time window (0-121 ms) presented more looks to the object in SVO than in OVS order, a second significant time window (1000-1606 ms) showed also more looks to the object in SVO than in OVS. The last significant time window (2545-3000 ms) showed more looks to the subject in SVO than in OVS order. When comparing the two conditions where the object referent was focused (Figure 3.7f), there were more looks to the object in SVO than OVS order in a very brief time window at the beginning (0-212 ms) and towards the very end of the segment (2364-2848 ms). In both comparisons, the differences appeared only sporadically and during relatively brief periods compared to the persistent effect of focusing displayed by the other comparisons.

To sum up, referents which were focus-marked in Segment 1 received more looks and therefore more attention than non-focus marked referents in Segment 2 as well, which can also be seen in Figure 3.6. Again, like in Segment 1, this implies that the focus-marked referent was more prominent / accessible than the non-focus marked referent.
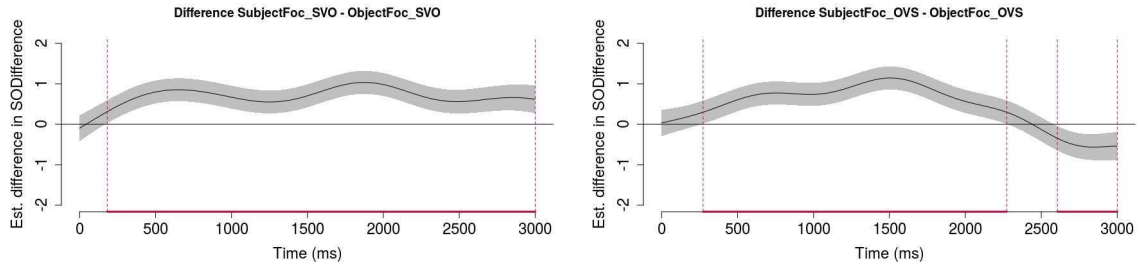
(a) Difference in looks between subject focus and object in SVO

(b) Difference in looks between subject focus and object focus for OVS
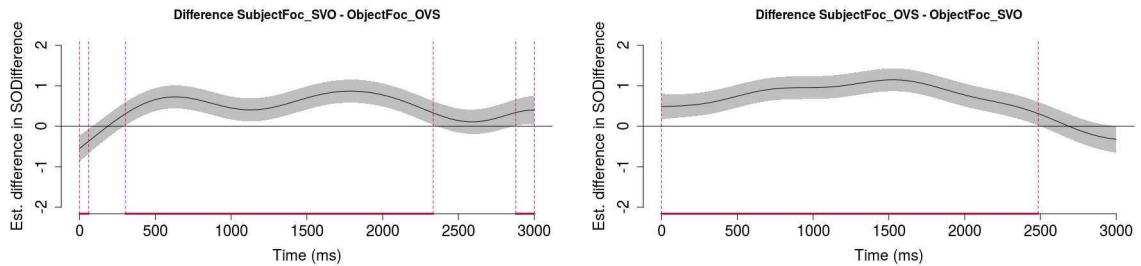
(c) Difference between subject focus SVO and object focus OVS

(d) Difference between subject focus OVS and subject focus SVO
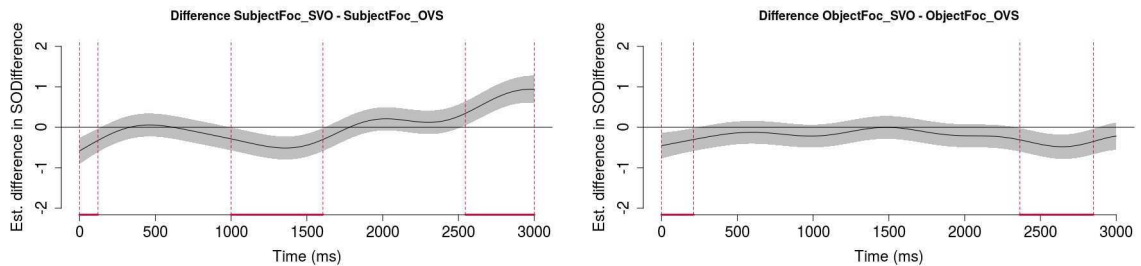
(e) Difference between subject focus in SVO and OVS

(f) Difference between object focus in SVO and OVS

Figure 3.7: Difference plots across all four conditions for the segment with the ambiguous pronoun. Shaded areas represent 95% confidence interval. Black line indicates zero effect.

Discussion

The offline results showed a robust preference for the pronoun *er* 'he' to be resolved towards the subject referent, confirming the subject preference (Bader & Portele, 2019b; Crawley & Stevenson, 1990; Frederiksen, 1981; Fukumura & van Gompel, 2015; Gordon & Chan, 1995; Kaiser, 2011a; Song & Fisher, 2005).

Since this effect was observed regardless of word order, and word order also did not affect referent selection (i.e. neither the subject nor the object referent were chosen more often when they were mentioned first), we can assume that non-

linguistic factors, i.e. first mention, play a negligible role in pronoun resolution in German (also see Schumacher et al., 2016, 2017). Instead, our results suggest that in German, a linguistic factor, i.e. grammatical role, is the primary determinant of pronoun interpretation. Furthermore, and contrary to expectation, the subject referent was the preferred referent even in the object focus condition, indicating that subjecthood was a stronger cue than focus during pronoun resolution.

For online processing, the Segment 1 (prosody and word order manipulation) revealed a robust effect of focusing showing that participants directed their attention towards the focused referent. When the focused referent was not in first-mentioned position, the shift to the focused referent started somewhat later. However, even with this small delay, the attention shifted to the focus marked referent happened relatively fast. Additionally, participants attended to the focus marked referent equally regardless of its grammatical role, suggesting that, as expected, prosodic focus marking was the driving cue for more attention towards the referent for this segment.

During the pronoun segment (Segment 2), overall attention towards the referents decreased compared to the previous segment, but was still higher for the focus marked referent than for the non-focus marked referent. As predicted, the focus-marked referent received more looks regardless of its grammatical role (also cf. Figure 3.6). This indicates that the effect of prosody from Segment 1 carried over to Segment 2, suggesting that in Segment 2 as well, focus-marked referents were more prominent than non-focus marked referent. However, this effect did not influence final interpretation of the pronoun (cf. Figure 3.3).

So, prosody, but not word order, can be an additional cue to boost attention towards a specific referent. Similar to Cowles et al.'s (2007) findings, focus enhanced referents' cognitive accessibility. In the object focus condition, this led to increased attention to the object referent. Thus, focusing the object referent hindered the pronoun's resolution towards the (preferred) subject referent to some extent. The focus effect was not strong enough to override effects of subjecthood, rather it seemed to have made reference interpretation somewhat less straightforward compared to the subject focus condition, which was reflected in the numeri-

92

cal, though statistically not significant, decrease in the subject choices in the offline data.

Regarding the effects of information structure, it is interesting that the results indicated focus effects in the gaze data, with no significant influence in the offline data, but also did not show a topic preference. Thus, the present experiment failed to replicate previous research finding a topic or "anti-focus" preference for subject pronouns (Colonna et al., 2012, 2015; De la Fuente & Hemforth, 2013; Patterson et al., 2017), as well as the observations on a general linguistic preference for topic continuity (Colonna et al., 2012; De la Fuente & Hemforth, 2013).

In summary, the experiment's results showed that both online processing and offline interpretation of the personal subject pronoun *er* 'he' were guided exclusively by linguistic factors, with the final interpretation showing a strong subject preference. The second linguistic factor, information structure, only influenced online processing, in the form of focus-marked referents receiving more attention than non-focus marked referents. In contrast, order of mention did not have a significant effect.

However, at this point it is not clear if these findings are particular to the pronoun *er* or whether they generally hold for (German) subject pronoun resolution. In particular, the form-specific approach to reference resolution Kaiser & Trueswell (2008) predicts that different pronouns will be sensitive to different factors. To test this, we conducted a follow-up experiment comparing two different subject pronouns, the personal pronoun *er* and the demonstrative pronoun *der*. Based on the results of the present study, we predicted that referent resolution for both pronouns would be guided by linguistic factors, but that the degree of sensitivity towards these factors would differ between the pronouns, as predicted by the form-specific account.

3.6   Experiment 2: referent selection for *der* and *er*

Experiment 2 investigated to what extent prosodic marking of information structure can influence the referent preference for the personal pronoun *er* and the

demonstrative pronoun *der*. Personal pronouns and demonstrative pronouns have been found to have different preferences with regards to the choice of antecedents. One way these preferences have been described is with regards to grammatical role (Diessel, 1999): while unstressed personal pronouns prefer the subject, demonstrative pronouns prefer the object. Referent preferences have also been explained using information structure (Bosch et al., 2007, 2003; Comrie, 1997): unstressed personal pronouns have been observed to prefer topical referents, whereas the non-topical referent is selected by the demonstrative pronoun. In previous studies, demonstrative pronouns have been found to be affected more by information structure than personal pronouns (e.g. Bader & Portele, 2019a; Ellert et al., 2011; Kaiser & Trueswell, 2008). Thus, we expected the demonstrative pronoun to show a clear preference for the object referent in the object focus condition, since the preferences for object and non-topic align in the same referent in this condition. In the subject focus condition, we do not expect the demonstrative pronoun's preference for a referent to be as obvious, since the preference for the non-topic should counteract the preference for the object.

3.6.1 Methods

Participants

We collected data online from 107 monolingual speakers of German (85 female, 21 male, 1 not reported; age range: 19-68, mean: 32, sd: 11.11). All participants reported normal hearing.

Materials

In Experiment 2, we used the same 40 dialogues as in Experiment 1, i.e. we applied the same prosody manipulation as in Experiment 1 (subject focus vs object focus; cf. Table 3.1). However, we kept only SVO order. Furthermore, we varied the pronoun at the beginning of Segment 2. In addition to the unaccented personal pronoun *er*, we included the unaccented demonstrative pronoun *der*. We again distributed the 40 experimental items across four lists in a Latin square design, and added the same 40 filler items as Experiment 1 to each list.

Procedure

The data was collected using Google Forms. Links to one of the four lists were distributed through mailing lists. In the beginning of the experiment, participants were instructed to listen to the dialogues and answer a question regarding the pronoun's referent following each dialogue.

3.6.2 Results

We performed generalized linear mixed-effects regression modeling (GLMER) using the lme4 package (version 1.1-23, Bates et al., 2015) in the software R (version 3.6.3, R Core Team, 2019) to analyze the participants' responses. The models included a binomial dependent variable coding whether the participant chose the subject or the object as the referent of the pronoun (*RefPreference*; excluding 17 responses choosing a distractor referent, i.e. 0.4% of the data, leaving 4263 data points for analysis). We added an interaction between focus (subject focus vs object focus) and pronoun (*er* vs *der*), and a by-participant random slope for pronoun as well as a random intercept for item. Opting again for a backward-fitting model comparison procedure, we excluded fixed factors one by one to see whether they contributed to the goodness of fit of the model. The interaction between focus and pronoun significantly contributed to model fit and was therefore retained.

As can be seen in Figure 3.8, the subject referent was more often chosen in the subject focus condition compared to the object focus condition. The best-fitting statistical model indicated that this difference was significant (see Table 3.5). With regards to the pronouns, *er* was the preferably resolved towards the subject referent, while participants selected the subject referent significantly less often for *der*. The subject preference for *er* was modulated by focus marking, i.e. the preference in the subject focus condition was more robust than in the object focus condition. The effect on the interpretation of the demonstrative pronoun was even larger, as indicated by the significant interaction between pronoun and focus marking. In the object focus condition, the preferred referent was the object. In the subject focus condition, however, the object referent was not preferred over the subject referent,

but instead, participants choose subject and object referents almost equally often.
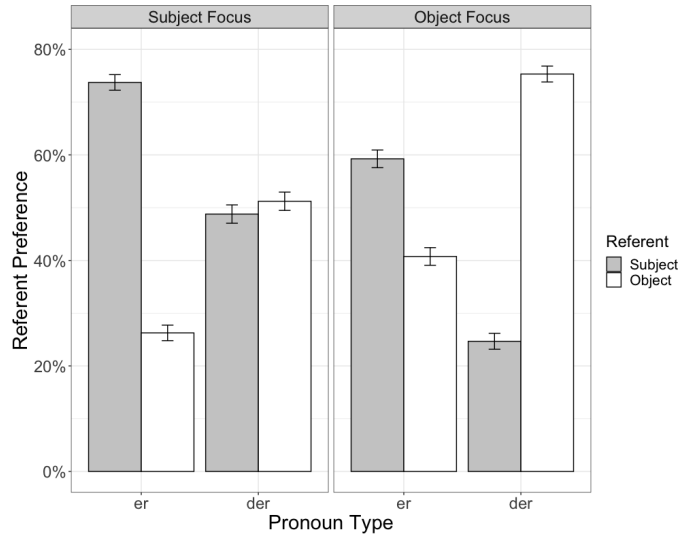
Figure 3.8: Referent choice for er and der by prosody, in percent, with error bars for standard error.

Table 3.5: Fixed effects for best fitting generalized linear mixed-effects model of referent choice in Experiment 2.

Formula: RefPreference ~ Focus * Pronoun + (1 + Pronoun | Participant) + (1 | Item)				
	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	1.4992	0.2376	6.311	2.77e-10
ConditionObject Focus	-0.8473	0.1130	-7.498	6.48e-14
PronounDER	-1.5550	0.1802	-8.629	< 2e-16
ConditionObject Focus:PronounDER	-0.6437	0.1593	-4.040	5.35e-05

We further analysed the interaction by comparing the multiple levels using the lsmeans package (version 2.30-0, Lenth, 2016). The multiple comparison revealed that there were significant differences in referent choice among all levels of comparison.

3.6.3 Discussion

The results of Experiment 2 confirmed the strong subject referent preference found in Experiment 1 for the personal pronoun. Prosodic focus marking influenced the subjecthood effect, but could not overwrite it. In line with previous findings on personal and demonstrative pronouns in German and Finnish (e.g. Bader & Portele, 2019a; Ellert, 2013; Kaiser & Trueswell, 2008; Schumacher et al., 2016), the

demonstrative pronoun was more sensitive to focused referents than the personal pronoun, as well as overall more likely to be resolved towards the object referent. The topic vs non-topic preference (Bosch et al., 2003) mentioned above holds to a certain degree for our results. While it correctly captures the preference for the non-topical referent found for demonstrative pronouns, the preference for the topical referent for the personal pronouns is not supported by our results. Like Experiment 1, Experiment 2 points to subjecthood being the main factor to drive referent preference for the personal pronoun.

Overall, the interpretation of both pronouns was guided clearly by linguistic factors, grammatical role and information structure, but differed regarding the pronoun's sensitivity to each factor. While the the personal pronoun preferred the subject referent, the demonstrative was affected by objecthood. Focus marking played a role for both pronominal forms, however as also depicted in Figure 3.8 it was a stronger cue for the demonstrative pronoun than for the personal pronoun. This experiment showed that that these two pronouns differ in their referent preferences, which is in line with the form-specific approach (Kaiser & Trueswell, 2008), but not with approaches in terms of prominence.

3.7   General Discussion

In this study, we addressed whether linguistic factors, i.e. subjecthood or information structure, or non-linguistic processes, i.e. first mention, would be decisive in pronoun resolution (e.g. Frederiksen, 1981; Gernsbacher et al., 1989). In particular, while the effect of information structure marking on pronoun resolution has been investigated before, it has most of the time been explored without the use of appropriate context, for example, to render explicit focus marking felicitous. Moreover, information structure marking has been generally used in the form of syntactic focus marking not taking prosody into account (e.g. Colonna et al., 2012, 2015; Järvikivi et al., 2014; Patterson et al., 2017). German, a language with a relatively free word order, allowed us to tease apart subjecthood from first mention and information structure effects. While changes in word order may evoke changes in in-

formation structure (e.g. Frey, 2005, 2006), using prosodic focus marking allowed us to also control referents' information status.

Experiment 1 strongly suggested that the resolution of the personal pronoun *er* is determined only by linguistic factors, i.e. subjecthood. Experiment 2 replicated the findings for *er* and revealed that the demonstrative pronoun *der* was influenced by objecthood as well as focus marking. Further, it showed that the demonstrative pronoun was more affected by focus marking than the personal pronoun (also see Kaiser & Trueswell, 2008, for Finnish). These results confirmed the previously found differences in linguistic preference for *er* and *der* (e.g. Ariel, 1990; Bader & Portele, 2019a; Bosch et al., 2007; Comrie, 1997; Diessel, 1999) and is in line with Kaiser & Trueswell's (2008) form-specific account.

Further, our findings indicated that prominence cannot simply be used as an umbrella term to explain why certain referents are preferred over others. In an attempt to substantiate the assumed prominence of pronoun referents, which has been criticized as circular (see Section 3.1), we analyzed looks to referents as a proxy for attention, which has explicitly been linked to prominence/accessibility/ salience in the literature (e.g. Brennan, 1995; Grosz & Sidner, 1986; Grosz et al., 1995; Gundel et al., 1993). The eye-tracking data in Experiment 1 gave us insight about what happens before the pronoun appears. In Segment 1 (prosody and word order manipulation), the focus marked referent obtained significantly more looks, i.e. more attention, than the given topic-marked referent regardless of its grammatical role or word order. The effect of focus marking carried over to the pronoun segment (Segment 2). Here as well, focus-marked referents (from Segment 1) received more looks than given topic-marked referents (from Segment 1). As mentioned before, we assume that more looks towards the focus marked referent equals more attention towards that referent which in turn should mean that the referent is more prominent/accessible (e.g. Gundel et al., 1993; von Heusinger & Schumacher, 2019).

Importantly, the offline data revealed that the observed attention towards the focus marked referent did not transfer to and determine the final interpretation: a preference for the subject referent was found in all conditions regardless of order-

of-mention or information structure. Thus, the idea that pronouns are resolved towards a certain referent because it is prominent/accessible/salient does not seem to hold. The gaze data of the first segment together with the offline results further imply that the subject referent is not selected as the pronoun's referent because it is prominent/accessible/salient – otherwise the focused object referent would have been the preferred referent with the object focus condition – but because it is the grammatical subject (and/or possibly agent, see discussion below).

There is a possible counter-argument to this reasoning, namely that the focused constituent may have consistently been the most prominent referent around the time when the pronoun was encountered, as well as during the preceding sentence, but maybe the relative prominence of referents still changed after that and before finally determining pronoun resolution. According to von Heusinger & Schumacher (2019) and Tomaszewicz-Özakın & Schumacher (2022), the computation of prominence profiles is dynamic (see also section 3.1). This would mean that computation would not end when encountering the pronoun. In other words, maybe we failed to substantiate the idea that prominence determines pronoun resolution because we looked at the wrong stretches of time?

According to Tomaszewicz-Özakın & Schumacher (2022), prominence profiles are fully computed at the predicate and available upon encountering the pronoun. That said, this would only reflect the prominence at the point in time during processing when the pronoun is encountered. Maybe the contrast between the offline results and participants' attention to the focused referent at that point could be explained by the idea that the computation of prominence profiles is not completed when encountering the pronoun, but that it carries on after? However, even considering participants' attention to the referents at the end of the pronoun sentence (the next time point when the prominence profile would be updated following von Heusinger & Schumacher (2019) and Tomaszewicz-Ö zakın & Schumacher (2022) and the last relevant point before the offline question), participants' attention is not consistent with the offline results of Experiment 1. In particular, the object focus conditions initially showed more looks to the object which decreased towards the end of the pronoun sentence. This resulted in both referents receiving an equal

amount of looks, i.e. attention, at the end of Segment 2 (cf. Figure 3.6). Thus, if we take more attention towards one referent as that referent being more prominent,[5] then prominence cannot be an explanation as to why the offline data still revealed the subject to be the preferred referent.

Since the preference for a certain referent does not follow from its salience, whether this is due to order-of-mention or focus, we argue that subjecthood itself is guiding resolution of the personal pronoun *er*. Particularly in pronoun resolution research, terms such as salience or prominence seem to have been applied to the grammatical subjects in an attempt to explain why they are the preferred referent for ambiguous personal pronouns. However, the present results indicate that this approach holds little if any explanatory power. Moreover, as indicated by the results of Experiment 2, this preference is particular to this subject pronoun. This is in line with previous research on Finnish pronouns and supports the form-specific approach to reference resolution, according to which different pronouns cannot be mapped onto a single salience hierarchy, but are each sensitive to different factors (Kaiser, 2005a,b; Kaiser & Trueswell, 2008). Moreover, as Kaiser & Trueswell explain, demonstrative pronouns do not mirror the use of personal pronouns, i.e., their preferences are not simply based on opposite factors (e.g. subject vs object role). This was well captured in the results of Experiment 2 where overall the personal pronoun preferred the subject referent. Yet focus still modulated choice of the preferred referent as the subject was chosen to a lesser degree when the object referent was focused. In contrast, the demonstrative pronoun showed the degree to which focus affected referent choice to be higher. When comparing the results of *der* in the subject focus condition to *er* in the object focus condition (see Figure 3.8), unlike *er*, *der* did not show a preference for a referent. Thus, the degree

[5]A devil's advocate might of course argue that we have failed to find effects of prominence / accessibility / salience because of this assumption. However, not only is the link between prominence / accessibility / salience and attention made explicitly in much of the literature on pronoun resolution (and beyond) that appeals to this concept (e.g. Brennan, 1995; Grosz & Sidner, 1986; Grosz et al., 1995; Gundel et al., 1993), the idea that the most prominent referent would *not* correspond to the referent attracting most attention also seems to run counter to the core observations and intuitions that linguists have tried to capture by using the terms accessibility, prominence or salience (cf. overviews and discussion in von Heusinger & Schumacher, 2019; Ladd & Arvaniti, 2023).

to which multiple factors influence referent choice is dependent on the pronominal form: personal pronouns exhibit a stronger link to subjecthood (grammatical role), whereas demonstrative pronouns exhibit a stronger link to focus (information structure). Crucially, both pronouns are sensitive to linguistic factors.

In the application of the form-specific approach to further languages, we would therefore also expect that each pronoun within a language would be affected by a different combination of linguistic factors. Differences may also appear between languages, so while personal subject pronouns prefer subject referents in Finnish, English and German, pro-drop languages that allow null subject pronouns such as Italian (Carminati, 2002), Spanish (Alonso-Ovalle, Fernández-Solera, Frazier & Charles, 2002; Gelormini-Lezama & Almor, 2011), or Japanese (Okuma, 2011) too show a preference for the subject referent but for the null form. So, the degree to which other factors exert an influence on pronominal forms, as well as the factors themselves could vary among languages. An extension of the form-based approach should therefore compare pronouns in a larger sample of languages, which could also lend further evidence to its main hypothesis, that preferences of different pronouns cannot be captured in terms of a simple hierarchy of prominence, accessibility or salience.

What this approach leaves unexplained so far is why it is subjecthood that the personal pronoun is sensitive to. In German, research has shown that in the interpretation of (temporarily) ambiguous sentences, there is a strong bias for a parse that assigns the subject role to the first entity encountered in the discourse (Mecklinger et al., 1995; Meng & Bader, 2000a,b; Schriefers et al., 1995). This means that in sentence processing, the subject is generally expected to appear before the object, unless e.g. context licenses object-first interpretations (see e.g. Weskott, 2003). This preference for the subject-first interpretation may also transfer to pronoun resolution where the personal pronoun is interpreted to be linked to the previous subject referent. When encountering a pronoun, the ambiguity does not lie with the grammatical role, since this is already encoded in the pronoun itself, but with which referent the pronoun needs to be connected. If the choice were between a subject and an object referent, following the subject-first account for ambiguity, the parser

would opt for the subject referent, unless information in the discourse disagrees with that interpretation e.g. verb semantics (e.g. Järvikivi et al., 2017). Similarly, Kaiser & Trueswell (2008) suggest that early subject looks in SVO conditions (regardless of pronoun type) stem from the expectation that the subject referent will be mentioned again which they see in line with predictions from Centering Theory (e.g. Grosz et al., 1995) that it would be easier to process a subject if it had been the preceding discourse-old subject. Easier processing in turn would result in the prediction of subjects to occur frequently.

However, while these suggestions further seem to support the idea that there is a subject bias in processing, they do not explain it. One possible avenue to pursue in search of an explanation, suggested to us by an anonymous reviewer, is that the subject bias stems from the fact that subjects are usually also agents, as discussed below. Interestingly, the gaze data in Experiment 1 suggests that the subject bias can be overruled during processing, but seems to still hold for the final interpretation.

Another question that then emerges is why this subject-bias only holds for the personal pronoun but not for the demonstrative pronoun. A simple answer in line with the form-specific approach would be that there is no need for two subject pronouns with the same referential function. If we look at *der* as not being the 'default' form for a subject pronoun, perhaps its use implies that we should not be opting for the 'default' interpretation, i.e. the subject referent. This would be in line with the complementary preferences (subject vs object, topic vs non-topic) that have been described for the two referential forms (e.g. Bosch et al., 2007; Diessel, 1999).

One question that remains is whether the subject preference for *er* is solely an interpretation preference for the subject or also a preference for the agent, since generally these two roles are confounded in the same referent, as mentioned above. In this study we only teased apart grammatical role from order of mention and thus the subjecthood effect might simultaneously be an effect of semantic role. Other studies that disentangled grammatical from semantic role Schumacher et al. (2016, 2017) have argued for agentivity rather than subjecthood to be the main

predictor for pronoun resolution (see e.g. Leslie, 1996, for an analysis of agency).

However, the results of those studies differed with regards to their robustness for default and non-default word orders (Schumacher et al., 2017), but also depending on whether assumed preferences for grammatical role and semantic role aligned (Schumacher et al., 2016, 2017). The preferred referent in the gaze data seemed not to be as clear with non-default word orders. This could be due to focus marking not coinciding with the object referent. For Schumacher et al. (2017), it is likely that the first-mention was interpreted as the topic (subject in SVO, object in OVS) and the second-mention was interpreted as the focus (object in SVO, subject in OVS). As our results from Experiment 2 showed, there was no referent preference for *der* when objecthood and focus did not align, which may also be the case for *der* in Schumacher et al. (2017). But for their study, the role of information structure was not clear since there was no context or prosody manipulation that would ensure a constant assignment of information structural roles across word orders. The non-default word order could have also caused broad focus or contrastive focus interpretations (cf. B ü ring, 1997; Fanselow, 2015; Frey, 2005, 2006; Jacobs, 1991) which may have influenced referent selection consequently. This could be the reason why referent preference in Schumacher et al. (2016) was not as clear with non-default word orders.

It also has to be noted that Schumacher et al.'s (2016; 2017) studies investigating the effects of semantic role use the conjunction *aber* 'but' right before the pronoun in their test items. In literature, the use of *but* has been considered to indicate not only contrast (Sæbø, 2019; Winterstein, 2012) but also topic change (Sæbø, 2019; Umbach, 2005). Thus, it is not clear to what extent this could have also influenced their pronoun results. As can be seen in Schumacher et al.'s (2016) third experiment, which did not use *aber* 'but', preference bias changed for *er* with experiencer dative verbs (chance level) when compared to experiment two in which *aber* was used. It is therefore likely that pronoun resolution was influenced by information structural change due to the use of *aber* 'but' to some degree. Nonetheless, our results are potentially consistent with Schumacher et al.'s (2016) results, showing a preference to resolve *er* towards the subject/agent and *der* towards the ob-

ject/patient referent.

To conclude, the results of our study provide evidence that linguistic factors determine pronoun resolution: the main factor for the personal pronoun was grammatical role, while the demonstrative pronoun was affected by a combination of grammatical role and information structure. Further, the preference for the subject referent for the personal pronoun *er* follows not from the subject being prominent, but from a general preference for the subject (and/or agent) in sentence processing, or perhaps the status of subjects (and/or agents) in languages in general (Polinsky, 2011).[6] We took the amount of looks to a referent as reflecting their prominence and showed that prominence was not the cause for the subject preference for *er*, because prominence in this sense would have deemed the object referent the preferred interpretation in the object focus condition in Experiment 1. Since this was not the case, prominence cannot be the explanation as to why subjects are the preferred referent for *er* (cf. also Polinsky, 2011). For the demonstrative pronoun *der*, we found a combination of objecthood and focus marking to influence referent selection. The results are not surprising as it seems intuitive for the two different pronominal forms to display different levels of sensitivity towards grammatical role and focus marking – otherwise what would be the use for another pronominal form? It is also in line with the form-specific approach of Kaiser & Trueswell (2008), which states that personal pronoun and demonstrative pronouns are sensitive to different factors. We can also assume that *er* and *der* will show different preferences with regards to agentivity, as has been found in Schumacher et al. (2016, 2017), however, the extent of sensitivity the two pronominal forms display is yet to be investigated. Further, this study has also shown that the special role that subjects (and/or agents) seem to hold for pronoun resolution, and sentence

---

[6]As (Polinsky, 2011, 324) notes: "In the study presented here, adult heritage speakers [...] reformatted the syntactic design of Russian in such a way that the only robust relativization position is that of the syntactic subject. This is consistent with the accessibility hierarchy (Keenan & Comrie, 1977), but the consistency is not an explanation; it simply upholds a well-established empirical generalization. These results further reflect the special status of subjects, which surfaces in other grammatical phenomena such as the interpretation of anaphors: As in relativization, anaphors are also more likely to select a subject antecedent than a grammatical function lower on the scale (see Polinsky & Kluender, 2007, for further discussion). The general principle seems clear: Subjects have a special status across languages. Yet the question remains as to why the preference for subjects exists and what makes them special."

processing more generally, needs to be further examined.

Chapter 4

Understanding L1 and L2 Reference Comprehension in Speech: Focusing Referents and Pronouns

Abstract

In this study we present data from two experiments investigating the effect of information structure, prosodic focus marking, on German L1 and L2 speakers' interpretation of pronouns. Experiment 1 tested L2 speakers' interpretation of personal and demonstrative subject pronouns. Experiment 2 examined L1 and L2 speakers' interpretation of unaccented and accented personal subject and object pronouns. The results of experiment 1 reveal that L2 speakers are sensitive to the different functionalities of the two subject pronouns. However, since grammatical role and focus marking influenced referential choice to similar degrees for both pronouns, this suggests that L2 speakers seem not to have acquired that the two pronominal forms are affected to different degrees by these linguistic factors. Experiment 2 shows L1 and L2 speakers to prefer the subject referent for both subject and object pronouns. Referent preference reversal is only observed with the accented subject pronoun in L1 speakers. Ultimately, this study emphasizes the varying levels of sensitivity for grammatical role and information structure observed not only for the different pronoun types but also among different speaker groups.

4.1   Introduction

For successful language comprehension, new information has to be stored and linked to already existing information. The use of a pronoun suggests a referent that is already known. Since pronouns themselves generally encode little information (e.g., person, number), linking the pronoun to a referent is necessary to retrieve more information for the comprehension process. Pronoun resolution, in turn, is sensitive to a wide variety of factors which can affect the likelihood of a certain referent being linked to a pronominal form, among them grammatical role (e.g. Alonso-Ovalle et al., 2002; Carminati, 2002; Crawley & Stevenson, 1990; Frederiksen, 1981; Fukumura & van Gompel, 2015; Gordon & Chan, 1995; Järvikivi et al., 2005; Kaiser, 2011a; Okuma, 2011; Song & Fisher, 2005) and information structure (e.g. Colonna et al., 2012, 2014, 2015; Ellert, 2013; de la Fuente & Hemforth, 2013; Xu, 2015), but the factors and their relative weights have been shown to differ between languages and individual pronominal forms (e.g. Bader & Portele, 2019a,b; Ellert, 2010; Ellert et al., 2011; Kaiser & Trueswell, 2008).

In this study we investigate the effects of prosodic focus marking on first (L1) and second language (L2) speakers' pronoun interpretation. Experiment 1 tests how focus marking on preceding subject and object referents influences referential choice for personal pronoun *er* and demonstrative pronoun *der* for L2 speakers of German. In experiment 2, we employ prosodic focus marking in the form of accents on subject and object pronouns themselves and examine its effect on referent selection for both L1 and L2 speakers. While the role of information structure of the preceding context on the interpretation of *er* and *der* has been investigated for both L1 and L2 speakers (Ellert, 2010; Ellert et al., 2011; Patterson et al., 2017; Wilson, 2009), these studies have not considered possible effects of prosodic focus marking. Thus, we want to expand the existing characterization of the interpretative preferences of *er* and *der* by disentangling effects of information structure and grammatical role for L2 speakers. With regards to stressed or accented pronouns, thus far, the effect on referent selection has not been examined in German L1 or L2 speakers for both subject and object pronouns (but see e.g. Mozuraitis &

Heller, 2017; Taylor et al., 2013, for English). Our aim is therefore to contribute to the pronoun and L2 literature by addressing this research gap.

4.1.1 Pronoun Resolution in L1 German

For German, the personal subject pronoun *er* ("he") usually refers to the preceding subject referent (e.g. Bader & Portele, 2019b; Bouma & Hopp, 2007; Colonna et al., 2012; Hert et al., n.d., Chapter 3 in this thesis). In addition to the personal pronoun, demonstrative pronouns e.g., *der* are used anaphorically as well. The two pronominal forms – personal and demonstrative pronouns – have been found to differ in their preference regarding choice of referents. The difference for the two pronominal forms has been described in terms of complementary preferences for grammatical role as well as for information structure. Unstressed personal pronouns would prefer the subject/topical referent, demonstratives are more likely to be linked to the object/non-topical referent (Bosch et al., 2007, 2003; Comrie, 1997; Diessel, 1999; Kaiser, 2011b). However, studies targeting the differences in pronominal preferences reveal that their referent preference cannot be captured by simple complementarity. The extent of sensitivity towards these factors varies among the pronominal forms, as demonstratives are affected more by information structure than personal pronouns, whereas personal pronouns are influenced by grammatical role to a greater degree than demonstrative pronouns (e.g. Bader & Portele, 2019a; Hert et al., n.d., Chapter 3 in this thesis; Kaiser, 2011c; Kaiser & Trueswell, 2008; Portele & Bader, 2016). In fact, Kaiser & Trueswell (2008) put forward the form-specific approach that claims that multiple factors (e.g., subjecthood, focus) play a role for linking a specific referent to a specific pronoun and that the degree of sensitivity to these factors varies for the the different pronominal forms (e.g. personal pronoun vs demonstrative pronoun, or overt pronouns vs null pronouns).

The resolution of subject pronouns has been researched much more intensively, whereas fewer studies have examined the resolution of object pronouns. Moreover, research on object pronouns mainly used parallel structures like those in (10). Some studies revealed that the object pronoun was preferably resolved to-

wards the object referent, i.e. *Ross* (e.g. Kehler et al., 2008; Smyth, 1994; Stevenson, Nelson & Stenning, 1995), whereas other studies did not find this preference (e.g. Crawley, Stevenson & Kleinman, 1990; Kehler et al., 2008).

(10)   Joey hugs Ross, and Monica punches him.

Importantly, these studies mostly focused on English and tested three types of parallelism. The first type describes parallelism with respect to the referent and the pronoun sharing the same grammatical role. In the sentence above, *Ross* and *him* are both objects. The second one defines parallelism in terms of position. Both *Ross* and *him* appear in sentence final position or as the second-mentioned entity. And lastly, structural parallelism describes the similarities between referent and pronoun sentences with respect to their sentence structures. In the example above, both sentences follow the S(ubject)-V(erb)-O(bject) order. The third type in particular has been found to affect sentence processing in English (e.g. Callahan, Shapiro & Love, 2010; Frazier, Taft, Roeper, Clifton & Ehrlich, 1984; Poirier, Walenski & Shapiro, 2012) but also in German (e.g. Knoeferle & Crocker, 2009; Weskott, 2003). Smyth (1994) and Stevenson et al. (1995) aimed to disentangle the different parallelism factors. Smyth (1994) found both grammatical role and structural parallelism to influence resolution. Stevenson et al. (1995) suggest not all factors to contribute to pronoun resolution, specifically, parallel position did not determine pronoun resolution outcomes. Similarly, Kehler et al. (2008) investigated differences in referent assignment between parallel coherence, (11a), and result coherence, (11b), structures for English with object (and subject) pronouns.

(11)   Samuel threatened Justin with a knife, and (Kehler et al., 2008, 12)

    a.   Erin blindfolded him.

    b.   Erin stopped him.

For parallel structures, they found the pronoun was preferably linked to the referent of the corresponding grammatical role, i.e. *Justin* in (11a). For result structures, participants preferred the referent of the opposite grammatical role, i.e. *Samuel* in (11b). However, neither of these studies tested structures in a non-canonical word order.

Using eye-tracking, Sauermann & Gagarina (2017) investigated the effects of word order and grammatical role parallelism during online pronoun processing in German, which allowed them to untangle the effects of positional and structural parallelism with regards to grammatical role. The gaze data showed a stable effect of grammatical role, i.e. the subject referent was fixated with subject pronouns, while the object referent was attended with object pronouns. This effect appeared irrespective of word order. However, recent research shows that online preferences in the visual world do not necessarily reflect the final interpretation (cf. Blything et al., 2021b; Hert et al., n.d., Chapter 3 in this thesis; Schumacher et al., 2016, 2017). Sauermann & Gagarina also note that the effect of grammatical role parallelism was not present in the last two time segments. They assume that this effect might decrease during later processing.

Following Sauermann & Gagarina's experimental design, Abashidze et al. (2022) examined the influence of grammatical and positional parallelism during online and offline resolution of subject and object pronouns in German. However, unlike Sauermann & Gagarina (2017), they only tested SVO order. Their gaze data revealed an initial preference towards the subject referent for subject and object pronouns which increased further in the subject pronoun condition. The offline results showed a preference for the subject referent with the subject pronoun, while for the object pronoun, referent choice was at chance-level. Similar to Sauermann & Gagarina, Abashidze et al. (2022) found grammatical parallelism to strongly affect online processing. Yet, unlike Sauermann & Gagarina, the gaze pattern for object pronouns showed a stronger subject referent preference. Abashidze et al. explain their results in terms of topicality. They see personal pronouns as a tool for topic continuation and participants would apply a topic bias during their resolution. Resolution of subject pronouns in topic position would be straightforward because grammatical role and topicality align. On the other hand, the resolution of object pronouns would not be as clear, since grammatical role and topicality would be disjoined, which might explain the chance-level performance on the offline interpretation.

Most studies on pronoun resolution use the unaccented/unstressed form of a

pronoun as the default intonation. According to Akmajian & Jackendoff (1970), accented or stressed pronouns receive a different interpretation than unaccented pronouns. In her Accessibility Theory, Ariel (1988, 1990, 2001) proposed that referential expressions differ in terms of their mental accessibility, where stressed pronouns would mark lower accessibility than unstressed pronouns. So, when it comes to linking the pronoun to a referent in the discourse, an unstressed pronoun would be resolved towards the (discourse) topic (cf. Bosch et al., 2007, 2003), whereas a stressed pronoun would take the nontopical referent (see also e.g. Givón, 1983; Gundel et al., 1993; Kameyama, 1999; Smyth, 1994, for similar accounts). This reversal pattern has also been attained in coherence-based accounts (e.g. Kehler, 2005). Here, accenting does not target the referent but rather the event. Linking a referent to the accented pronoun that differs from the default referent indicates that events unfold unexpectedly (see also Wagner & Watson, 2010, for an overview on prosodic prominence and predictability).

Taylor et al. (2013) tested the influence of accented and unaccented object pronouns in English (and Spanish) in two experiments. In the first experiment (see (12) for an example of their items), they found a bias for the object referent with unaccented object pronouns. The preference was reversed when the pronoun was accented in English (or via pronoun doubling in Spanish).

(12)   Rachel texted Monica and Ross called her/HER.

The second experiment included a strong verb bias for one referent (e.g., *Joey arrested Ross$_1$ and Monica released him$_1$*). The results showed a change in preference for the referent with accented pronouns (double pronouns in Spanish), although, unlike in the first experiment, preference was only reduced but not reversed. Thus, Taylor et al.'s findings suggest that the reverse resolution pattern for accented pronouns holds, but other factors, such as verb bias, can affect the likelihood of a pronoun being resolved towards a certain referent and also influence the degree of reversal. Taylor et al. (2013) therefore assume that the reverse interpretation is blocked when the alternative referent is not plausible. In line with Taylor et al.'s hypothesis, Gillingham (2013) argues that an accented pronoun does not always

lead to a switch in referential links as its felicity depends on how unlikely the default referent is in a given context. If the context would suggest one referent to be more likely than the other (for an unaccented pronoun), then the use of an accented pronoun would be felicitous to refer to the less likely referent. In contrast, Mozuraitis & Heller (2017) show that an absence of reversal pattern can be noticed even when the alternative referent is plausible. They propose that the interpretation of accenting requires alternatives Rooth (1985) that are readily available in the discourse context. In the case of accented pronouns, it involves contrasting the interpretation of the verb phrase that contains the pronoun with an alternative. If the discourse context provides an alternative that changes the default interpretation of the pronoun, a reverse in interpretation can occur. According to Mozuraitis & Heller (2017, p. 101), example (13a) – a parallel structure – but not (13b) – a result structure – would be an alternative to *sent HIM a postcard*.

(13) The animals went on a school exchange across the globe.

  a. Pig mailed Elephant a souvenir.

  b. Pig gave Elephant his address.

Then, Bear sent him/HIM a postcard.

Mozuraitis & Heller further note that thus far all experimental work on accented pronouns found reversed interpretation patterns only in parallel structures (e.g. Venditti, Stone, Nanda & Tepper, 2002). Moreover, most research has focused on English which, unlike German, does not use demonstrative pronouns anaphorically. So, perhaps, since there is an additional pronoun in German that differs in its referential function from the personal pronoun, referential function will also differ with respect to accented pronouns when compared to English. And while Zerbian & Böttcher (2019) found accented pronouns to occur in spontaneous speech in German L1 and bilingual speakers, they did not investigate the pronoun's referential function. Nonetheless, support for this idea may come from Gargiulo, Tronnier & Bernardini's (2019) production experiment with L1 Italian and L1 Swedish speakers. They found the use of pronouns produced with a "higher degree of prominence" (i.e., relatively longer, louder and with a wider tonal excursion of F0)

differed for the two languages. In Italian accented pronouns were linked to the subject referent, whereas in Swedish they were linked to the object referent. So, if Italian and Swedish differ in their use of accented pronouns, the same could apply for German and English.

4.1.2 Pronoun Resolution in L2

In research on L2 acquisition, the focus often lies on linguistic phenomena which differ in the source language and the target language. With regards to pronoun resolution, especially pro-drop languages have been contrasted with languages that generally only allow overt pronouns. Comparing different pronominal systems, research investigates whether these differences affect pronoun resolution in L2 and whether L2 speakers can perform like L1 speakers.

Polio (1995) conducted a study to examine how null pronouns are acquired in Chinese, a pro-drop language. Native speakers of Chinese were contrasted to two L2 groups, with English (not pro-drop) and Japanese (pro-drop) as their L1. The results showed that both L2 groups used fewer null pronouns compared to L1 speakers in semi-spontaneous productions. Regarding L2 speakers' overuse of overt pronouns, L1 transfer can be excluded as a possible explanation for their different behaviour, since even Japanese speakers did not benefit from their L1 licensing null pronouns similarly to Chinese. While L2 speakers' use of null pronouns increased with proficiency, it did not seem to explain L1 and L2 differences either.

For Italian, a pro-drop language, Belletti et al. (2007) found that English near-native L2 speakers of Italian were significantly more likely to use overt pronouns in both spontaneous and elicited production tasks, as well as more likely to link overt pronouns to the subject of the main clause than Italian native speakers. The authors assume that L2 speakers' L1 – which is a non-pro-drop language – affects the use and interpretation of pronouns in their L2. They consider that L2 speakers have more options available for pronoun interpretation since they can access two grammatical systems, which makes different grammatical options (i.e., pronominal functions) more accessible to them than to monolingual speakers. Similar re-

sults were put forward in Tsimpli, Sorace, Heycock & Filiaci's study (2004). The authors showed that native Italian speakers under attrition whose L2 was English used more overt subject pronouns than the native control group.

Roberts et al. (2008) investigated if advanced L2 learners of Dutch would be affected by their L1 in their offline and online resolution of subject pronouns. For L1s, the authors contrasted a pro-drop language, Turkish, with a non-pro-drop language, German. The results for the offline data showed that the L1 German group performed like the Dutch native speakers preferring to link the (intra-sentential) pronoun to the subject referent, whereas the L1 Turkish group did not show a preference for either referent. Interestingly, for the online data the two L2 groups patterned together. Unlike the Dutch native speakers, neither L2 group preferred the (intra-sentential) subject referent during online processing. Even though the two L2 groups had the same online processing pattern, their final interpretations diverged. The authors conclude that L1 influences emerge at the level of pragmatics to resolve the syntactic ambiguity. This means that L2 speakers would take their L1 into consideration when difficulties in L2 processing would arise.

Schimke & Colonna (2016) tested pronoun interpretation of L1 Turkish learners of French, alongside French native speakers (in French) and Turkish native speakers (in Turkish), manipulating semantic roles and order of mention. Their findings revealed that when interpreting ambiguous overt subject pronouns in French, native French speakers were influenced by order of mention and semantics, i.e., preferring the semantic patient in initial position. Native Turkish speakers primarily relied on subjecthood instead, when interpreting null subjects in Turkish translation equivalents. L2 learners performed similar to native French speakers in that they preferred the first-mentioned patient referent. Additionally, topicality seemed to have affected their referent choice as well. The authors conclude that the L2 learners might rely on discourse-level information to a greater extent than L1 speakers when interpreting pronouns.

Okuma (2011) investigated the effects of explicit topic and subject marking (via particles) in L1 English learners of Japanese in comparison to a native Japanese control group. The results for null subject pronouns showed that native speak-

ers preferred the subject over the object referent regardless of topic marking. In contrast, L2 speakers opted more often for the subject referent than the object referent in OS order while in SO order there was no preference. With overt pronouns, native speakers did not have a preferred referent, whereas L2 learners linked the subject pronoun more often to the object referent, regardless of topic marking or word order. In sum, grammatical role influenced comprehension of overt and null pronouns in both L1 and L2 speakers to a greater extent than topicality. However, while subjecthood was the deciding factor in native speakers for the resolution of null subject pronouns, L2 speakers did not show the same behavior. In contrast, L2 speakers showed a preference with the overt pronoun where native speakers did not, namely objecthood being the determining factor for its resolution.

Lozano (2018) explored pronoun resolution in L2 Greek speakers of Spanish with varying levels of proficiency within different discourse contexts. Spanish and Greek both use a pronominal system with null and overt pronouns. However, the L2 speakers did not benefit from language similarities, and while higher proficiency resulted in a more native-like performance, even speakers of high proficiency differed from native speakers. The author puts forward that the disparity between L1 and L2 speakers would stem from L2 learners preferring being redundant rather than being ambiguous and therefore accepting overt pronouns more often than native speakers in topic-continuity contexts. This would mean that the use of overt pronouns is less ambiguous than the use of null pronouns. However, the study only tested acceptability of overt and null pronouns, but not their interpretation.

Abashidze et al. (2022) tested the role of grammatical and positional parallelism on subject and object pronoun resolution in German using eye-tracking as well as an offline interpretation task. They contrasted L2 speakers (L1 Georgian) with native German speakers. In the gaze data, they found L2 speakers to attend more to the subject referent than the object referent after a subject pronoun, which corresponds to L1 speakers' gaze pattern. For object pronouns, L2 speakers fixated the object referent more than L1 speakers. For the offline results, L2 speakers showed the same tendency as L1 speakers, namely, selecting the subject referent more of-

ten for the subject pronoun than for the object pronoun. However, L2 speakers preferred the object referent with the object pronoun, whereas L1 speakers did not show a preference for either referent. Abashidze et al. conclude that L2 speakers may have difficulties employing information structural cues and hence rely only on grammatical parallelism during pronoun resolution. Note, however, that their study did not manipulate information structure directly.

Wilson (2009) tested order of mention and grammatical effects on the resolution of German personal and demonstrative pronouns in L1 and L2 speakers. The results showed L2 speakers to prefer the first-mention – topical – referent with the personal pronoun, while L1 speakers showed no preference. For the demonstrative pronoun, L2 speakers had no preference, whereas L1 speakers linked the pronoun to the second mention – non-topical referent.

Ellert et al. (2011) investigated the effect of topicality on the resolution of personal and demonstrative pronouns in L2 German. Dutch is similar to German with regards to the referent preferences of personal and demonstrative pronouns (e.g. Bosch et al., 2007; Kaiser, 2011c). The L2 speakers linked both pronouns to the topical (first-mentioned) referent, which is different to the L1 preference to resolve the personal pronoun towards the topical (first-mentioned) referent and the demonstrative towards the non-topical (second-mentioned) referent. However, even though L2 speakers showed a topic preference for both pronouns, the preference was stronger for personal than for demonstrative pronouns. This was observed in their online eye-tracking data as well as in the offline comprehension questionnaire. These results suggest that the referential function of the individual pronouns played a decisive role in the final interpretation. Ellert (2010) further suggests that proficiency is a crucial factor in L2 resolution of personal and demonstrative pronouns. She observes that less proficient learners would use both pronominal forms for the same function (i.e., linked to topical referents), whereas highly proficient L2 learners would differentiate distinct functions for the personal pronoun (i.e., linked to topics) and for demonstrative pronouns (i.e., linked to non-topics).

Patterson et al. (2017) conducted three experiments to examine how focus af-

fects the resolution of within-sentence pronouns in both native and non-native German speakers, as well as native Russian speakers. Focus was established through the use of cleft constructions and focus-sensitive particles. The results indicated a distinct contrast between native and non-native speakers that could not be attributed to the L1 influence. Specifically, native speakers of German and Russian were less likely to link a pronoun with a referent in focus (via cleft) when compared to a non-focused referent in the same position. On the other hand, non-native speakers did not display this effect, but rather tended to resolve a pronoun towards referents appearing with a focus-sensitive particle. Thus, L2 speakers showed sensitivity to focus marking, but it had the opposite effect on their pronoun interpretation when compared to native speakers. This in turn would exclude possible L1 influences.

In sum, the studies presented in this section show that native and non-native speakers can differ in their choices with respect to pronominal referents. Despite proficiency playing an important role for L2 speakers to achieve a more native-like performance (e.g. Ellert, 2010; Lozano, 2018; Polio, 1995), research with highly proficient L2 speakers found the difference in referent choices between native and non-native speakers to persist (e.g. Belletti et al., 2007). Non-native speakers' L1 did not have a direct effect on performance in the L2. Even when L1 and L2 correspond in their licensing of pronominal use (e.g. both languages have different pronominal forms), L2 speakers still deviated from L1 speakers (e.g. Ellert et al., 2011; Lozano, 2018; Polio, 1995; Roberts et al., 2008). However, language influence in the sense of dominance/proficiency cannot completely be disregarded as Tsimpli et al. (2004) have shown even native speakers under attrition behave more like L2 than L1 speakers. This is also indicated in Roberts et al.'s (2008) study where L2 speakers of different L1s showed similar online processing patterns, but deviated in their final interpretation. So, perhaps if discourse shifts L1 and L2 attention to the same referent, L1 and L2 speakers' performance in pronoun interpretation will be more similar.

Regarding information structure, Schimke & Colonna (2016) suggested that L2 speakers would make use of it to a higher degree than L1 speakers, whereas

Abashidze et al. (2022) argued L2 speakers to have difficulties using information structure cues. A number of studies have also directly investigated information structural effects in L2 processing (e.g. Ellert et al., 2011; Okuma, 2011; Patterson et al., 2017). The studies demonstrated that L2 speakers were sensitive to changes in information structure, yet their referent choice was affected differently when compared to L1 speakers.

4.1.3 The Role of Prosodic Focus Marking in L1 and L2 Speakers

Intonation can be used as an indication of information structure. For instance, it can mark whether an element of a sentence has been introduced in the previous discourse (Schwarzschild, 1999), whether that element is new (i.e., update of the common ground, Lambrecht, 1994) or whether that element indicates the relevance of alternatives (Rooth, 1985, 1992). In German, focus is associated with a falling accent (H*(+L)) while post-focal given information is deaccented (e.g. Baumann, 2006; Bü ring, 1997; Féry, 1993). Topics, on the other hand, are connected to rising accents (L*+H), especially when contrastive (e.g. Braun, 2006; Bü ring, 1997; Féry, 1993; Repp & Drenhaus, 2015). Focus is acoustically marked with a wider pitch range, an increased intensity, and increased duration compared to other speech elements that are not in focus (Féry & Kügler, 2008).

The literature suggests that native speakers are able to identify and integrate prosodic information to build information structure in real time (Heim & Alter, 2006; Wang, Wang & Qadir, 2011). An ERP experiment in German (Hruska & Alter, 2013) revealed an increased N400 response at words that were expected to carry a focus pitch accent but did not, implying that prosody and expectations about focus may immediately influence processing.

In various studies (Chen et al., 2007; Ito et al., 2014; Ito & Speer, 2008; Sedivy et al., 1999), contrastive focus marking has been revealed to trigger anticipatory eye-movements, e.g. hearing *blue ball* followed by *GREY* raises expectations that the upcoming noun will also be *ball* (Ito et al., 2014; Ito & Speer, 2008), which in turn can support target search. Similarly, it has been observed that in H*L (focus) conditions, the initial higher proportion of looks directed towards the competitor

decreases earlier compared to L*H (non-focus) conditions (Chen et al., 2007; Sedivy et al., 1999). This shows that native listeners can make predictions about the upcoming referent in real time using prosodic cues.

For L2 learners' use of L2 prosodic cues, some studies suggest that learners have difficulties producing and perceiving prosodic cues, particularly if these differ from prosodic cues in their L1 (Mennen & De Leeuw, 2014). Akker & Cutler (2003) found L2 Dutch learners of English were not able to map pitch accent to semantic information as effectively as native speakers of English, even though the use of prosodic cues for information structure is similar in Dutch and English (Chen & Lai, 2011, also see). In contrast, Takahashi, Kao, Baek, Yeung, Hwang & Broselow (2018) found reaction times to be shorter for sentences with felicitous contrastive pitch accent as compared to infelicitous use of contrastive pitch for L2 Chinese learners of English. The authors assumed that the effective use of English contrastive prosodic cues in L2 speakers stems from similarities in pitch for prosodic focus in English and Mandarin Chinese. ERP studies (Reichle, 2010; Reichle & Birdsong, 2014) revealed that L2 proficiency can affect the online perception of information structure. Unlike low-proficiency L2 English learners of French, high-proficiency learners showed a native-like negative response for contrastive focus. Perdomo & Kaan (2021) looked into the effects of proficiency and working memory on L2 information structure processing. They found that while L2 speakers used prosodic information to build information structure during listening, neither proficiency nor working memory influenced L2 speakers' use of contrastive pitch accent to predict or process the following noun phrase.

The evidence thus far suggests that L2 speakers show sensitivity to modulations of L2 prosody, but may experience difficulties in effectively using the prosodic information for the subsequent discourse. As to the role of proficiency and distinctions in L1 and L2, these factors seem to not have a clear effect on L2 performance.

4.2 Current study

The aim of the current study is to test the influence of information structure on L1 and L2 pronoun resolution in German. More specifically, we want to find out how L1 and L2 speakers' interpretation of ambiguous pronouns is affected by prosodic marking. In addition, we consider effects of parallelism, particularly grammatical role and position parallelism (see elaboration below). In experiment 1, we investigate whether L2 speakers' referential choice can be aided by using focus on either of the two possible referents. That is, can their referent preference be biased towards one referent if that referent is explicitly marked with focus via prosody in the discourse context (see section 4.3.1 for an example)? Further, we want to examine whether L2 speakers are sensitive to the different referential functions of the subject pronouns *er* and *der*. Previous L2 research suggests that L2 speakers differ from L1 speakers in their resolution of the demonstrative pronoun *der* but are more alike for resolution of the personal pronoun *er* (e.g. Ellert et al., 2011). Therefore, we assume referential choice to be more alike to L1 speakers with *er*, whereas for the demonstrative we predict L2 speakers to perform close to chance level. With respect to information structure effects, it seems that L2 speakers show sensitivity, yet it does not lead to a native-like performance for pronoun resolution (Ellert et al., 2011; Patterson et al., 2017; Wilson, 2009). If L2 speakers are guided by focus marking rather than grammatical role, we expect the focused referents to be chosen more often than the topical referents.

In experiment 2, we test L1 and L2 speakers' referential choice for the subject pronoun *er* and object pronoun *ihn* (see section 4.4.1 for full example and all pronoun conditions). We manipulated whether pronouns were unaccented or accented, as in (14), to analyze if that would affect referent selection and if L1 speakers would differ from L2 speakers.

(14) *Der Arzt* bringt *den Koch* mit einer Clownsnase zum Lachen, als
      er/ER die Musikerin mit der Kamera filmt. (*subject pronoun*)

      '*The doctor* makes *the cook* laugh with a clown's nose, when
      he/HE filmed the musician with the camera.'

As discussed in section 4.1.1, previous research found unaccented personal subject pronouns to be preferably resolved towards the preceding subject referents (e.g. Abashidze et al., 2022; Bader & Portele, 2019b; Bouma & Hopp, 2007; Colonna et al., 2012; Hert et al., n.d., Chapter 3 in this thesis), whereas for the object pronoun, a preference for object referents (Sauermann & Gagarina, 2017) or topical and object referents (Abashidze et al., 2022) was observed for L1 German speakers. Therefore, for L1 speakers, we expect a clear preference to link the unaccented subject pronoun to the subject referent. If the interpretation of the unaccented object pronoun is indeed affected to the same extend by multiple factors as suggested by Abashidze et al. (2022), then we expect no preference for either referent. With respect to accented pronouns, based on previous findings, whether or not the accent would lead to a reversal in referential choices has been suggested to depend on the one hand on the probability of the context favouring the default interpretation for the unaccented pronoun (cf. Gillingham, 2013; Taylor et al., 2013). On the other hand Mozuraitis & Heller (2017), argue it depends on alternatives being explicitly available in the previous discourse, which would be the case for our experimental items. If reversal of the referent preference is dependent on how plausible the alternative referent is, then preference of the accented pronouns may not be as clear, as it could be different depending on the item. If reversal is indeed only observed in the parallel discourse type, we predict the accented subject pronoun to show a preference for the object referent. However, since we assume no preference for the unaccented object pronoun, that means there cannot be a preference for the accented pronoun either. In any case, both accounts assume that pronoun type (subject or object pronoun) does not influence the reversal effect. Therefore, if the reversal effect of accented pronouns is present, we expect to observe it for both pronouns.

Turning to L2 speakers, considering the research presented in section 4.1.2, we assume referent choice to be similar to that of L1 speakers with unaccented pronouns, but the preference may not be as distinct as with L1 speakers. Regarding resolution of accented pronouns, as Gargiulo et al. (2019) revealed prosodic prominence of pronouns to have a similar effect on L1 Italian and L1 Swedish, namely

favouring the less predictable referent, we conclude L2 speakers to react in a similar manner to accented pronouns as L1 speakers.

In addition to the accent, we included an order of mention manipulation for the object pronoun *ihn*. This was done so that we could target possible effects of parallel position (cf. Abashidze et al., 2022; Sauermann & Gagarina, 2017). Note that while Abashidze et al. (2022) and Sauermann & Gagarina (2017) have looked into the effect of parallel position as well, they actually did not manipulate order of mention of the pronoun. As Abashidze et al. (2022) also indicate, moving the object pronoun to the sentence initial position would emphasize its topicality. We therefore may see a difference in preference for the object pronoun when it is in first mention position compared to second mention position. Assuming that there is a bias for topical referents with pronouns, the preference for the topical referent may be more pronounced when the object pronoun is in first mention position compared to second mention position.

Lastly, in relation to previous work on L2 pronoun resolution mentioned in section 4.1.2, we also considered the role of proficiency and included a measure of vocabulary knowledge. Following previous findings (e.g. Lozano, 2018), we assume that L2 speakers' performance correlates to their level of proficiency. This means the higher their score on the proficiency measure, the closer their performance on referent selection should be to the native speakers'. Moreover, we expect proficiency to be of greater importance for pronoun resolution than L2 speakers' L1, considering that even native speakers under attrition performed more similar to L2 speakers than to native controls Tsimpli et al. (2004).

4.3 Experiment 1: Interpretation of *er* and *der* in L2 speakers

Experiment 1 investigated to what extent prosodic marking of information structure can affect referent preference for the personal pronoun *er* and the demonstrative pronoun *der* in L2 speakers. Design and materials were identical to those used with L1 speakers in experiment 2 reported in Hert et al. (n.d., Chapter 3 in this

thesis).

4.3.1   Method

Participants

80 participants with various L1s (see Table 4.1) completed the experiment via Prolific. For the analysis, we excluded 15 participants based on their high error rate with filler items (more than 60% incorrect), resulting in 65 participants (age range: 19 - 62, mean: 33, sd: 12.4). The study was approved by the Research Ethics Board 2 of the University of Alberta (study ID Pro00105075).

Table 4.1: L1s and number of speakers for experiment 1.

L1	N
English	12
Polish	10
Dutch	5
Greek, Hungarian, Portuguese, Spanish	4 (each)
French, Italian, Russian	3 (each)
Slovenian	2
Bosnian, Bulgarian, Croatian, Dutch/French, English/Irish, Latvian, Lithuanian, Mandarin, Romanian, Turkish, Ukrainian	1 (each)

Materials

We designed a comprehension task where participants listened to short dialogues. Two experimental factors were manipulated: (i) whether the dialogue contained an unaccented personal pronoun *er* or an unaccented demonstrative pronoun *der*; (ii) whether the possible referents (subject, object) for the pronouns were focused (see Table 4.2 for an example). Information structure was manipulated in a twofold way, coupling changes in the context with prosodic focus marking. A total of four conditions were tested and we used ten sentences per condition, which resulted in 40 experimental dialogues.

The dialogues were recorded using Shure SM10A headset microphones in a sound attenuated booth by two native speakers of German, one female and one male. The female speaker (a prosody researcher) recorded all the introduction and

critical sentences (A-turns in Table 4.2) for the experimental items while for the filler items it was the male speaker. The speaker uniformly produced a single falling accent on the focused constituent, while the rest of the sentence remained unaccented (represented as H* L-% in GToBI notation Grice et al., 2005, 2017).

Table 4.2: Example dialogue with critical manipulation in all four conditions. Prosodic focus marking in italics, unaccented ambiguous pronoun in bold. Contexts were identical for all conditions except where indicated with slashes and condition names in brackets. Note that information structure was manipulated prosodically in the critical sentence, as well as in the preceding context.

German	English translation
A: Als ich letzten Mittwoch in der Arztpraxis war, habe ich den Angler, den Gangster, den Bäcker und den Käsehändler gesehen. Sie waren für einen Erste Hilfe Kurs da. Als erstes hat jemand den Käsehändler behandelt (subject focus). / Als erstes hat der Gangster jemanden behandelt (object focus).	When I was at the medical office last Wednesday, I saw the angler, the gangster, the baker, and the cheesemonger. They were there for a first aid course. Firstly, someone treated the cheesemonger (subject focus). / Firstly, the gangster treated someone (object focus).
B: Und wer hat den Käsehändler behandelt? (subject focus) / Und wen hat der Gangster behandelt? (object focus)	B: And who treated the cheesemonger? (subject focus) / And whom did the gangster treat? (object focus)
A: *Der Gangster* hat den Käsehändler behandelt (subject focus), / Der Gangster hat *den Käsehändler* behandelt (object focus), und zwar mit einem Pflaster. **Er/Der** fand die Farbe des Pflasters total schön.	A: *The gangster* treated the cheesemonger (subject focus), / The gangster treated *the cheesemonger* (object focus), namely with a band-aid. **He** really liked the color of the band-aid.
B: Gut!	B: Good!
Q: Wer fand die Farbe schön?	Q: Who liked the colour?

The experimental items were distributed across four lists in a Latin square design. Additionally, we constructed 40 filler items which were the same across all lists. These fillers also contained four possible referents, but unlike the experimental items we did not include any ambiguous pronouns.

Procedure

The experiment was created with the *jsPsych* framework (version 7.2.1, de Leeuw, 2015), which is used for carrying out online experiments. The participants were given a brief written explanation of the tasks they were about to complete. First, participants filled out a questionnaire about their language background. Additionally, we included the German LexTALE (Lemhöfer & Broersma, 2012) as a measure of L2 speakers' vocabulary knowledge (see Table 4.3). Afterwards, a screen with instructions appeared, asking participants to carefully listen to the dialogues. They were also given the chance to check their speakers'/headphones' volume before starting the task.

While listening to the dialogues, participants saw the names of the four mentioned referents on the screen. Following each dialogue, they saw a question on the screen probing to which of the two target referents, subject or object, the pronoun referred (see last row in Table 4.2). We also included the other two referents as possible responses to ensure that participants paid attention during the experiment. Participants gave their answer by clicking on one of the names on the screen. The positions of the referents' names on the screen were randomized for each list. Halfway through the experiment participants were given a break.

4.3.2 Results

We performed generalized linear mixed-effects regression modeling (GLMER) using the lme4 package (version 1.1-23, Bates et al., 2015) in the software R (version 4.0.2, R Core Team, 2020) to analyze the participants' responses. The models included a binomial dependent variable coding whether the participant chose the subject or the object as the referent of the pronoun *RefPref*. We therefore excluded 111 responses choosing a distractor referent, i.e. 4.27% of the data, leaving 2489 data points for analysis. We added an interaction among *Condition* (subject focus vs object focus), *Pronoun* (*er* vs *der*), and *LexTALE* (centered). *L1_type* (a binary variable coding whether participants' L1 was a pro-drop language) was also included as fixed effects. Moreover, the model contained a by-*Participant* random slope for

Table 4.3: LexTALE scores (raw) for L2 speakers, including range, mean, and standard deviation.

range	mean	sd
11.25-81.25	44.003	13.736

Pronoun as well as a random intercept for *Item*. Opting for a backward-fitting model comparison procedure, we excluded fixed factors one by one to see whether they contributed to the goodness of fit of the model. The initial model did not converge, therefore we tested the 3-way interaction separately from *L1_type*. Neither *L1_type* nor interactions between pronoun, condition and LexTALE affected the model's fit significantly, and were therefore excluded from the final model shown in Table 4.4.
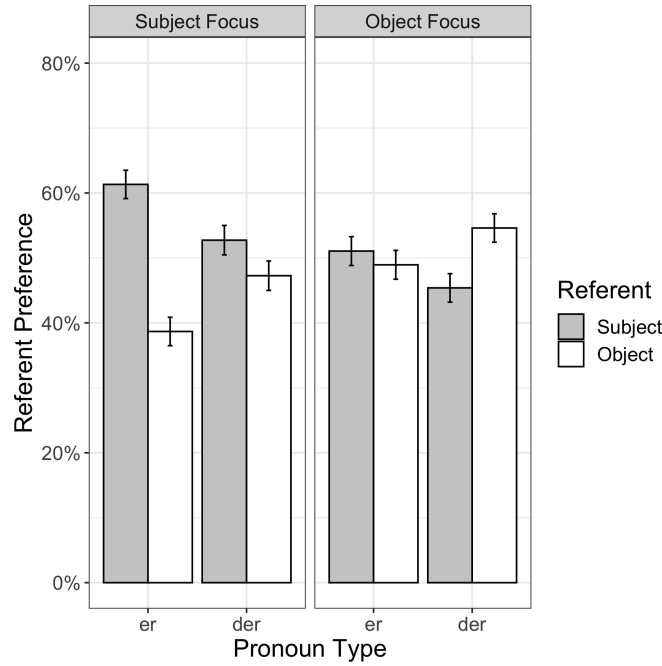
Figure 4.1: Referent choice for er and der by prosody, with error bars for standard error.

Table 4.4: Fixed effects for best fitting generalized linear mixed-effects model of referent choice for er and der.

Formula: RefPref ~ Condition + Pronoun + LexTALE + (1 + Pronoun − Participant) + (1 − item)				
	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	0.50665	0.13437	3.771	0.000163
Condition Object Focus	-0.42321	0.13642	-3.102	0.001921
Pronoun der	-0.33149	0.13787	-2.404	0.016202
LexTALE	0.16999	0.07325	2.321	0.020301

As can be seen in Figure 4.1, the subject referent was chosen less often in the object focus condition overall compared to the subject focus condition (intercept in Table 4.4). With regards to the pronouns, *er* was preferably resolved towards the subject referent, while participants selected the subject referent significantly less often for *der*. Moreover, the effect *LexTALE* influenced referent choice significantly. Higher scores resulted in an increase of the subject referent preference.

In addition, we compared the subject preference for the two pronouns in both prosody conditions to chance level using one-sample Wilcoxon signed rank test and found the difference to be significant only for *er* in subject focus and *der* in object focus conditions (for both $p<0.05$).

4.3.3 Discussion

Overall, the L2 speakers were sensitive to focus marking, as it affected their referent choice. Figure 4.1 depicts a preference for the subject referent in the subject focus condition with the personal pronoun *er*, and a preference for the object referent in the object focus condition for the demonstrative pronoun *der*. These two preferences were not as pronounced as for L1 speakers (cf. Hert et al., n.d., Chapter 3 in this thesis). Nonetheless, these results show that L2 speakers were sensitive to the different degrees to which multiple factors influence referent choice. Similar to L1 speakers, there was no preference for either referent with *der* in the subject focus condition for the L2 group. However, unlike L1 speakers who still preferred the subject referent, L2 speakers did not show a referent preference for *er* in the object focus condition. This suggests that L2 speakers do not show a strong preference for either *er* or *der*, but only in combination with focus marking did a preference for one referent emerge. More specifically, L2 speakers show some degree of sensitivity to the different preferences found in L1 speakers for the two different referential forms: *er* is more likely to be resolved towards the subject referent, *der* is more often linked to the object referent.

4.4 Experiment 2: Interpretation of *er* and *ihn* in L1 and L2 speakers

Experiment 2 investigated the effects of unaccented and accented personal subject pronoun *er* and personal object pronoun *ihn*, as well as the effect of order of mention for *ihn*, on referential choice in L1 and L2 speakers.

4.4.1 Methods

Participants

A total of 249 participants were recruited via Prolific as well as from the University of Kaiserslautern. After excluding German bilingual participants and participants who scored less than 60% correct on the filler items, data from a total of 220 L1 (n=113) and L2 (n=107) participants were analyzed (age range: 18-69, mean: 30, sd: 10.47). L2 participants varied in their L1s (see Table 4.5; pro-drop: n=94). Prolific participants received monetary compensation (£9/h), participants from the University of Kaiserslautern received a 10€ gift card as compensation for their participation. The study was approved by the Research Ethics Board 2 of the University of Alberta (study ID Pro00105075).

Table 4.5: L1s and number of L2 speakers for experiment 2.

L1	N
Polish	27
Spanish	22
English, Italian	7
Hungarian, Slovenian	6 (each)
Greek, Turkish	5 (each)
Russian	4
Czech, Dutch, Portuguese, Romanian, Ukrainian	2 (each)
Danish, English/French, French, Hungarian, Luxembourgish/English, Portuguese/Roman, Slovakian, Vietnamese	1 (each)

Materials

For the experimental items, we created 42 mini stories containing either a subject or an object pronoun; for an example see (15). The first sentence introduced a female referent first, followed by two male referents. The second sentence contained additional information about the first sentence, but did not include any of the referents. The third sentence was made up of a main clause, which repeated the two male referents, one as subject the other as object referent (in italics), and a sub- ordinate clause which contained an ambiguous pronoun (in bold), either subject (see 15a) or object pronoun either in first- or second-mention position (see 15b and 15c). Additionally, the pronouns were either unaccented or accented (marked with upper-case letters in (15)).

(15)   Die Musikerin, der Arzt und der Koch veranstalten gemeinsam eine Feier. Alle amüsieren sich gut. *Der Arzt* bringt *den Koch* mit einer Clownsnase zum Lachen, als

'The musician (female), the doctor (male) and the cook (male) host a party together. Everyone is enjoying themselves. *The doctor* makes *the cook* laugh with a clown's nose, when'

a. **er/ER** die Musikerin mit der Kamera filmt. (*subject pronoun*)

'**he/HE** filmed the musician with the camera.'

b. **ihn/IHN** die Musikerin mit der Kamera filmt. (*object pronoun first*)

'the musician filmed **him/HIM** with camera.'

c. die Musikerin **ihn/IHN** mit der Kamera filmt. (*object pronoun second*)

'the musician filmed **him/HIM** with a camera.'

The mini stories were recorded using a Shure SM10A headset microphone in a sound attenuated booth by a native speaker of German. The speaker (a prosody

researcher) uniformly produced a contrastive rising accent on the accented pronoun, which would be represented as L*H following Féry (1993) and as L+H* in GToBI notation Grice et al. (2005, 2017). Accented pronouns were approximately four times longer in duration than unaccented pronouns. Figure 4.2 shows an example for a pronoun in unaccented (Figure 4.2a) and accented (Figure 4.2b) conditions.

(a) er unaccented.   (b) er accented
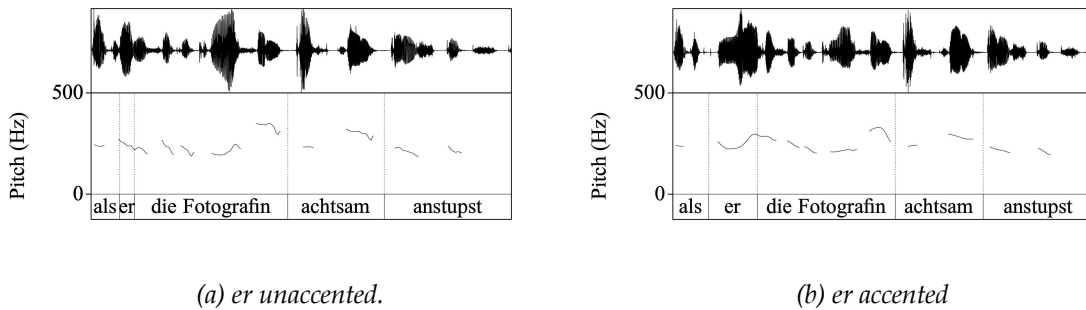
Figure 4.2: Prosodic contours for unaccented and accented pronouns.

Additionally, we constructed 28 filler items which were the same across all lists. These fillers also contained 3 referents, but unlike the experimental items, they had two female and one male referent and the pronoun was not ambiguous.

Procedure

The experiment was created using the *jsPsych* framework (version 7.2.1, de Leeuw, 2015). As in experiment 1, participants carried out a questionnaire and LexTALE (see Table 4.6), which were followed by the main task. While listening to the stories, participants saw the names of the three mentioned referents on the screen. Following each story, they saw a question on the screen asking them to choose the subject, object or another male referent that was not mentioned in the story as the referent of the pronoun. For subject pronouns, we asked a subject question (e.g. Who filmed the musician?), for the object pronouns, we asked an object question (e.g. Who did the musician film?). The positions of the referents' names on the screen were randomized for each list.

Table 4.6: LexTALE scores (raw) for L1 and L2 speakers, including range, mean, and standard deviation.

	range	mean	sd
L1	68.75 - 100	89.45	6.98
L2 non-pro-drop	56.25 - 87.5	71.25	8.09
L2 pro-drop	47.5 - 97.5	69.99	11.74

4.4.2 Results

We performed generalized linear mixed-effects regression modeling (GLMER) using the lme4 package (version 1.1-23, Bates et al., 2015) in the software R (version 4.0.2, R Core Team, 2020) to analyze the participants' responses. The models included a binomial dependent variable coding whether the participant chose the subject or the object as the referent of the pronoun, *ReferentChoice*. We excluded answers that selected a distractor (0.84%), which resulted in a total of 9162 observations. *Language_Type* (L1, L2 pro-drop or L2 non-pro-drop), *Pronoun* (*er*, *1_ihn*, or *2_ihn*), *Prosody* (unaccented or accented) and the centered *LexTALE* score were included as fixed effects as well as a 4-way interaction term. For random effects, we included a by-*Participant* random slope for *Pronoun* and a random intercept for *Item* (random slopes for *Prosody* lead to convergence issues). *LexTALE* did not contribute to any significant interactions and only remained as a fixed effect. The final model in Table 4.7 shows an effect for *LexTALE*, as well as for the interaction of *Language_Type*, *Pronoun*, and *Prosody*.

Both *Language_Type* and *Pronoun* consist of three levels. In order to see where the significant difference between the different levels of factors involved in the interaction are, we ran a pairwise comparison using *lsmeans* (version 2.30-0 Lenth, 2016). Table 4.8 shows the significant differences among these factor level combinations.

As can be seen in Figure 4.3, overall, there was a preference for the subject referent for both pronouns with all groups. However, within L1 speakers there was a difference for *er*; the subject referent was selected significantly less often for accented than unaccented *er* (see Table 4.8). The subject referent was also selected significantly less often for accented *er* compared to accented first- and

Table 4.7: Fixed effects for best fitting generalized linear mixed-effects model of referent choice for er and der.

Formula: ReferentChoice ~ Language_Type * Pronoun * Prosody + LexTALE + (1 + Pronoun − Participant) + (1 − item)				
	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	0.96358	0.14773	6.522	< 0.0001
Language_Type L2_drop	-0.50772	0.21146	-2.401	0.016348
Language_Type L2_non-pro-drop	-0.35764	0.36574	-0.978	0.328147
Pronoun 2_ihn	0.08350	0.15525	0.538	0.590679
Pronoun er	-1.03808	0.16266	-6.382	< 0.0001
Prosody unaccented	0.44781	0.15488	2.891	0.003836
LexTALE	0.32162	0.08557	3.759	0.000171
Language_Type L2_drop: Pronoun 2_ihn	-0.25043	0.17660	-1.418	0.156181
Language_Type L2_non-pro-drop: Pronoun 2_ihn	-0.18378	0.35033	-0.525	0.599876
Language_Type L2_drop: Pronoun er	0.85739	0.19652	4.363	< 0.0001
Language_Type L2_non-pro-drop: Pronoun er	1.14694	0.40144	2.857	0.004275
Language_Type L2_drop: Prosody unaccented	-0.49874	0.17578	-2.837	0.004549
Language_Type L2_non-pro-drop: Prosody unaccented	-0.13172	0.35106	-0.375	0.707498
Pronoun 2_ihn: Prosody unaccented	-0.45367	0.21637	-2.097	0.036015
Pronoun er: Prosody unaccented	0.88552	0.21650	4.090	< 0.0001
Language_Type L2_drop: Pronoun 2_ihn: Prosody unaccented	0.62080	0.24567	2.527	0.011506
Language_Type L2_non-pro-drop: Pronoun 2_ihn: Prosody unaccented	0.49260	0.49099	1.003	0.315727
Language_Type L2_drop: Pronoun er: Prosody unaccented	-0.60016	0.24686	-2.431	0.015051
Language_Type L2_non-pro-drop: Pronoun er: Prosody unaccented	-0.92894	0.49904	-1.861	0.062682

*Table 4.8: Multilevel comparison of the interaction term Language * Pronoun * Prosody of the generalized mixed-effects model. Positive estimates indicate a higher bias for the subject referent for the left factor level combination in the pair.*

contrasts	estimate	SE	df	z.ratio	p.value
L1 er accented - L1 er unaccented	-1.333324	0.151	Inf	-8.811	< 0.0001
L1 1_ihn accented - L1 er accented	1.038079	0.163	Inf	6.382	< 0.0001
L1 2_ihn accented - L1 er accented	1.121583	0.158	Inf	7.119	< 0.0001
L1 1_ihn unaccented - L2_drop 1_ihn unaccented	1.006453	0.216	Inf	4.668	0.0004
L1 2_ihn accented - L2_drop 2_ihn accented	0.758143	0.201	Inf	3.768	0.0192

second-mention *ihn*. While there were no significant differences between the two L2 speaker groups, for unaccented first-mention *ihn* and accented second-mention *ihn*, L1 speakers chose the subject referent more often than L2 speakers with pro-drop L1s. Finally, the positive effect of LexTALE in Table 4.7 indicates that the overall subject preference was higher for higher-proficiency speakers.

4.4.3 Discussion

The results of this experiment revealed differences, but also similarities between L1 and L2 speakers' choices for the subject pronoun *er* and the object pronoun *ihn*. First, L1 speakers showed an overall subject referent preference for all pronouns, with the exception of accented *er*. L2 speakers showed the same preference for the subject referent. However, unlike the L1 group, L2 speakers' choices for
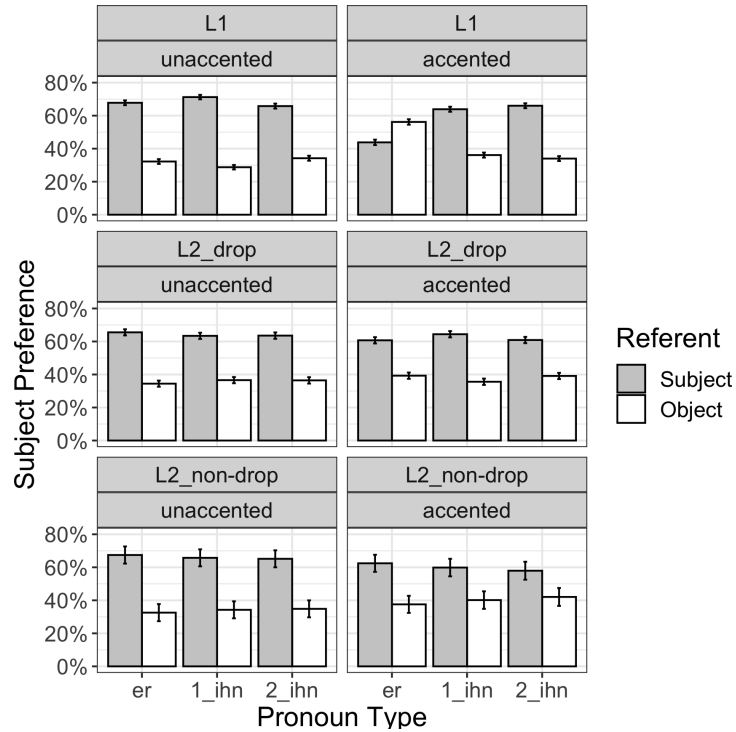
Figure 4.3: Referent choice for L1 and L2 speakers, accented and unaccented for er and ihn, with error bars for standard error.

the subject referent did not decrease when *er* was accented. Proficiency affected referent choice in that the subject referent was selected more often overall when participant's LexTale was higher, like in experiment 1. Moreover, type of L1 affected referent selection in L2 speakers. L2 speakers whose L1 was a non-pro-drop language performed more native-like than L2 speakers with pro-drop L1s as significant differences between L1 speakers and L2 speakers only occurred for the pro-drop group (see Table 4.8).

4.5 General Discussion

The two experiments in this study investigated the effects of prosodic focus marking in the preceding context on referent choice in L1 and L2 speakers of the subject pronouns *er* (personal) and *der* (demonstrative), as well as the effects of accents on the subject pronoun *er* and the object pronoun *ihn* (personal). In the first experiment, we asked if L2 speakers' referent selection would be affected if the referents

in the previous discourse were prosodically marked with focus. L2 speakers preferred the focused subject referent with the personal pronoun *er*, and chose the focused object referent more often for the demonstrative *der*. There was no preference for either referent when *der* was used in the subject focus condition, and when *er* was used in object focus condition. The results differ from L1 speakers' referent selection as the subject referent was preferred for *er* in both focus conditions, al- though focus modulated referent selection to some extent. For the demonstrative pronoun, L1 speakers preferred the object referent in the object focus condition but showed not preference in the subject focus condition (cf. Hert et al., n.d., Chapter 3 in this thesis). This suggests that grammatical role and focus marking impact the two referential forms to different degrees in L1. In our L2 speakers, neither grammatical role nor information structure seemed to have a clear effect on re- ferent selection by themselves. Nonetheless, it appears that L2 speakers acquired the different referential functions of *er* and *der* to some degree. If referent selection depended solely on the focus effect, we would expect to see a preference for the object referent in the object focus condition with *er* and a preference for the subject referent in the subject focus condition with *der*. Yet, for these two conditions pre- ference was at chance-level. And while the performance on *der* in the subject focus condition actually reflects L1 speakers' performance, this is not the case for *er* in object focus.

Another notable result is the effect of proficiency (LexTALE) on L2 speakers' performance. Higher proficiency levels led to the subject referent being selected more often for both pronouns. This contradicts the findings in previous research (e.g. Ellert et al., 2011; Lozano, 2018; Polio, 1995), that higher proficiency in L2 speakers would lead to a more native-like pattern for pronoun resolution. While there was a general tendency to select the subject more often as the referent for *er* than *der*, it could be that the overall proficiency of our L2 speakers was too low (mean: 44.003) for a native-like pattern to emerge.

In experiment 2, we investigated referential choice for unaccented and accented subject pronoun *er* and object pronoun *ihn* for L1 and L2 speakers. The results showed an overall preference to link the subject and the object pronoun to the

subject referent. L2 speakers' preference was reduced in comparison to L1 speakers, although pairwise comparisons were only significant compared to L2 speakers from the pro-drop group. Another difference emerged at the accented condition for *er*. Here, L1 speakers' preference decreased to chance level. This was not the case for L2 speakers who still selected the subject referent more often than the object referent.

We will first discuss these results for L1 speakers. In respect of unaccented pronouns, we addressed in section 4.1.1 that referent preference for subject and object pronouns has been mostly described in terms of grammatical role parallelism (Kehler et al., 2008; Sauermann & Gagarina, 2017; Smyth, 1994; Stevenson et al., 1995) or more recently (Abashidze et al., 2022), in terms of a combination of grammatical role parallelism and topic bias. Abashidze et al. argue that this combination would lead L1 speakers to not show a referent preference for the object pronoun, since objecthood and topicality do not align with the default information structure. Moreover, Abashidze et al. (2022) claim that moving the object pronoun into first mention position would enhance the topicality effect. In contrast, our L1 speakers preferred the subject referent regardless of pronoun type. Therefore, our results do not support the parallelism effects. Further, we did not find evidence for an enhanced topicality effect when the object pronoun was in first mention position compared to second mention position as there was no difference in preference for these two conditions. So, how can we explain the different results in our study compared to Abashidze et al. (2022) and Sauermann & Gagarina (2017)? One possible explanation stems from the experimental items. Abashidze et al. and Sauermann & Gagarina used contexts containing only two referents in their experimental items, whereas our items included a third referent that differed in gender. As already noted by Patterson & Schumacher (2021), a limitation in many exper- iments on pronoun resolution is that the contexts are limited to two potential re- ferents. However, including a third referent can give a more complex information structure (see Mozuraitis & Heller, 2017, for discussion on alternatives) which in turn may give a more detailed description of how sensitive different pronouns are to the factors involved in pronoun resolution (e.g. grammatical role, information

structure). The difference in referent preference for the object pronoun observed in our study may be the result of a more complex information structure.

Turning now to accented pronouns, previous research has established Gillingham (2013); Mozuraitis & Heller (2017); Taylor et al. (2013) that accented pronouns do not always lead to a reversal in referent preference, as had been initially claimed (e.g. Akmajian & Jackendoff, 1970). Indeed, our findings for the object pronoun do not show a switch in referent preference for the accented pronouns. However, we do find a switch for the accented subject pronoun. Neither Gillingham and Taylor et al.'s nor Mozuraitis & Heller's explanations as to why reversal patterns may not be triggered can account for the difference between subject and object pronouns. They all assume pronoun type not to be of importance for whether or not the discourse context gives rise to preference reversal. Again, a possible reason for our results could lie in the discourse contexts being more complex as they include three referents. Therefore, our results may actually point out a more detailed picture of the pronoun types' sensitivity to factors such as information structure and subjecthood (cf. Kaiser & Trueswell, 2008). This would imply that object pronouns are more restricted regarding the degree to which some of these factors can affect referential links than subject pronouns. In our study, object pronouns may show a greater preference for subjecthood than information structure, since changes in information structure (i.e., unaccented vs accented pronouns) did not influence referential choice. This would also explain why there was no effect of 'enhanced topicality' Abashidze et al. (2022) when the object pronoun was first mention compared to when it was second mention.

As for the L2 speakers, except for the accented subject pronoun condition, their performance was native-like in that they also preferred the subject referent for both pronouns. However, when compared to L1 speakers, L2 speakers linked the pronouns to the subject referent to a lesser extent, though the difference only reached significance for some factor level combinations (also cf. Abashidze et al., 2022). As for the reversal effect in non-native, it is not clear why it did not emerge. Our findings do not indicate a clear difference between pro-drop and non-pro-drop L1s, since both L2 groups showed the same subject bias (they also had very similar

mean LexTALE scores, cf. Table 4.6, see discussion below). Note that the group size for the L2 non-pro-drop speakers is a lot smaller than the other L2 or L1 groups, which may have obscured some possible differences. Regardless, L2 speakers' strategy during pronoun resolution seems to rely overall more on subjecthood than L1 speakers'.

Returning to the effect of proficiency, in the second experiment as well, we find higher LexTALE scores to lead to an increase in the subject referent choices for both pronouns. Even though the effect was the same in both experiments, the second experiment seems to show a more native-like performance than the first experiment. The difference in performance may be attributed to the overall level of proficiency being higher in experiment 2 than in experiment 1. We therefore assume that proficiency plays an important role in achieving a more native-like pattern of pronoun interpretation. In fact, if we take proficiency to reflect language dominance, then this could also explain why speakers under L1 attrition perform less native-like Tsimpli et al. (2004) when compared to native speakers. We propose that language dominance affects how native-like speakers' performance is during pronoun resolution. Unexpectedly, we also find L2 speakers' L1 to influence referential preference, but only in experiment 2. This may be due to different pronoun types being tested in the two experiments. While experiment 1 contrasted two types of subject pronouns, experiment 2 compared two pronouns of different grammatical roles. Another factor that could have influenced the difference in performance may be the experimental items. In experiment 1, the dialogues narrow down to two referents, whereas experiment 2 uses three referents. As discussed above, this could have led to a more complex information structure in experiment 2, which made otherwise subtle differences of pronominal sensitivity to different linguistic factors visible.

In conclusion, the two experiments have shown that L2 speakers are sensitive to focus marking during pronoun resolution. The study further supports the importance of proficiency to achieve native patterns in pronoun interpretation. For stressed pronouns, we provide evidence that the cause for the reversal pattern cannot be simply explained by whether information present in the discourse con-

text makes a switch in preference probable (Gillingham, 2013; Mozuraitis & Heller, 2017; Taylor et al., 2013), but that other factors must be included. For instance, the effect of grammatical role needs to be further investigated as we have only found referent preference to switch with the accented subject pronoun, but not with the accented object pronoun. Further, we did not find evidence for position of the object pronoun to affect its degree of topicality as has been claimed by Abashidze et al. (2022). Our study's limitations give rise to further investigate pronoun resolution in not only in L2 but also in L1 speakers with respect to testing effects of number of possible referents in the discourse, the influence of L2 speakers' L1 on pronoun resolution, as well as including different types of pronouns in one experiment.

Chapter 5

General Discussion and Conclusion

Throughout this dissertation, we investigated how information structure effects vary across different pronouns as well as different participant groups. The first (1), more general, question we aimed to address was how information structure affects processing of referents and pronouns in discourse contexts. Particularly, we wanted to know whether information structure manipulations on referents triggers changes in the processing of pronouns and ultimately changes in referent selection. Conversely, do information structure manipulations on the pronoun affect referent selection? Second (2), we wanted to address the question whether information structure would be processed differently in the different participant groups under investigation, specifically, whether children and L2 speakers would benefit in their referent selection – i.e. become more like adult L1 speakers – when changes in information structure shift attention to a particular referent. Finally, we examined whether the different types of pronouns show different levels of sensitivity to the effect of information structure.

While information structure has been manipulated in previous research, the findings were not always converging (cf. Colonna et al., 2012, 2014; Ellert et al., 2011; Järvikivi et al., 2005, 2014; Kaiser & Trueswell, 2008; Patterson et al., 2017). The goal of the present dissertation was to contribute to existing research on pronoun resolution by investigating effects of information structure in discourse contexts in German. The studies in this dissertation were designed to cover a variety of pronoun types, information structure manipulations, and participant groups. By doing so, this enabled us to receive a more detailed picture of the interplay of

these three factors.

5.1   Summary of Results

In study 1, we examined the information structural effects of *it*-clefts during online processing and final interpretation of reflexives and object pronouns in German children and adults. Based on findings of Spenader et al. (2009) and van Rij et al. (2016), we wanted to find out whether changes in information structure would shift attention to the focused referent and whether that would affect subsequent processing of the object pronoun and reflexive. The gaze data showed that the segment preceding the pronominal segment was processed similarly in children and adults in the sense that both groups paid more attention to the referent when it was clefted. In the pronominal segment, the groups diverged. Children paid overall more attention to the object pronoun referent than to the reflexive referent, whereas adults fixated the reflexive referent more than the object pronoun referent. For the last segment, gaze patterns of children and adults converged again. When the sentence matched the displayed picture, both groups fixated the intended referent of the pronominal element more than the competitor referent, e.g. when the sentence contained a pronoun, the pronoun referent received more attention than the reflexive referent. When the sentence did not match the displayed picture, the gaze pattern was mirrored, i.e. there was more attention towards the competitor than the intended referent. For the offline data, there was no effect of information structure.

In study 2, we investigated effects of subjecthood, information structure and order of mention for subject pronoun resolution in German adults. Information structure was operationalized as prosodic focus marking on either the subject or the object referent preceding the pronoun. The gaze data in experiment 1 showed focused referents received more looks than non-focused referents regardless of grammatical role or order of mention. The effect continued in the pronoun region and disappeared towards the end of the segment resulting in no preference for either referent. For referent selection in experiment 1 and 2, the personal pronoun

was preferably linked to the subject referent regardless of order of mention. Focus marking reduced the subject preference, but despite that, it was still the preferred referent. On the other hand, the demonstrative pronoun was preferably resolved towards the focused object referent. When the subject referent was focused, there was no preference for either referent.

In study 3, the first experiment investigated grammatical role and information structure effects on L2 speakers' referent selection for personal and demonstrative subject pronouns. The results revealed that a preference for either pronoun emerged only in combination with focus marking. For the personal pronoun, when the subject referent was focused it was the preferred referent. Similarly, for the demonstrative pronoun, the object referent was selected more often when it was focused in the preceding discourse context. There was no preference for either referent in the other two conditions. In experiment 2, we examined referent selection for unaccented and accented personal subject and object pronouns in L1 and L2 German. For L1 speakers, except for the accented subject pronoun, there was a preference for the subject referent. The object referent was selected more often with the accented subject pronoun. For L2 speakers, the subject preference was present in all conditions.

5.2 Discussion of Research Question

The first research question attended the effects of information structure during processing of referents and pronouns in discourse. More specifically, we asked whether information structure manipulations on referents would lead to changes in the processing of pronouns and/or referent selection. In study 1, gaze pattern for the segment with the two referents showed that there were more looks to a referent when it marked with focus by the *it*-cleft. Thus, marking the referent with focus resulted in more attention to that referent for both children and adults. It is only in the second segment, containing the pronominal elements, that children's gaze pattern diverged from adults. While children pay overall more attention to the object pronoun referent than the reflexive referent, adults attend the reflexive

referent to a great extent than the pronoun referent. We proposed that these differences in gaze patterns stem from differences in how children and adults process the preceding discourse. While children's processing is local, that is they only take the context immediately preceding the pronouns into account (cf. Trueswell et al., 1999), adults' processing is more global. We assume that children's limited working memory is the cause for their local discourse processing (e.g. Trueswell et al., 1999; Vogelzang, Guasti, van Rijn & Hendriks, 2021a). However, ultimately the differences between conditions as well as between groups during online processing did not affect offline choices, indicating that focus guides participants' attention during online processing, but final interpretation is restricted by the pronoun itself. Similarly, in the first experiment of study 2, prosodic focus marking led to the focused referent receiving more attention than the non-focused referent regardless of grammatical role. This effect was still present during processing of the subject pronoun. Yet, in the referent choices, we find – even though the preference is reduced with object referent focus – a preference for the subject referent. In the second experiment, referent selection for the demonstrative pronoun is guided by objecthood and focus marking. Moreover, results show that focus marking affects the demonstrative pronoun to a greater extent than the personal pronoun. These results led us to suggest that relative prominence of a referent cannot be the crucial factor why a pronoun is linked to a certain referent. If prominence was the decisive factor in pronoun resolution, we would expect to see a preference for the object referent in the object focus condition since it received more attention than the subject referent. However, that was not the case. With the personal pronoun, the prefe- rence for the subject referent was stronger suggesting a special status for the subject role and/or other factors, such as agentivity and topicality, that often coincide with the grammatical subject. And while these factors have been said to make a referent prominent (e.g. Arnold & Griffin, 2007; Arnold et al., 2013; Blything et al., 2021b; Ferreira et al., 2005; Foraker & McElree, 2007; Gundel et al., 1993; Kehler, 2002), it is not clear why exactly one factor would make a referent relatively more prominent than another factor. On the contrary, focus clearly affected referent prominence in the sense that it drew more attention to the referent that was focus-marked. For the

L2 group in study 3, experiment 1, preferences only emerge when focus is involved with the preferred grammatical role of the pronoun (i.e., subject role for personal pronouns, object role for demonstrative pronoun). This means focus affected referent selection in such a way that preferences similar to native speakers' arose. The final experiment also showed mixed results regarding information structure effects on referent selection. While generally the subject referent was preferred for both pronouns and both L1 and L2 speakers, this preference shifted to object referents with the accented subject pronoun but only for the L1 group. While it is surprising for this shift to only be present with subject pronouns, it may actually point out that the sensitivity for information structure effects differs for subject and object pronouns.

To sum up, regarding the effects of information structure on referent and pronoun processing, we found effects of focus marking leading to more attention on the focused referent and the effect continued to be present at the pronoun. Nonetheless, the presence of the effect during online processing does not necessarily mean that it will transfer to the final interpretation. This is especially the case in study 1 where there are no differences in participants' performance between the no cleft and cleft conditions. In the first experiment of study 2, we see a similar picture: focus shifts attention during online processing, but it does not shift the subject preference. For the second experiment of study 2, focus shifts referent preference but only for the demonstrative pronoun. Likewise, for the last two experiments, focus marking led L2 speakers towards a preference for one referent, whereas preference for accented pronouns was only switched for the subject pronoun in L1 speakers. The absence of the reversal effect in L2 speakers may have been due to an overall stronger preference for subjecthood or L1 influence. So, the effect that information structure can have during referent and pronoun processing but also final interpretation depends on pronoun type (cf. e.g. Bader & Portele, 2019a,b; Ellert, 2013; Ellert et al., 2011; Kaiser & Trueswell, 2008), the manipulation used, as well as the participant group under investigation (cf. e.g. Ellert et al., 2011; Järvikivi et al., 2014; Patterson et al., 2017). This means for the different pronouns that the demonstrative subject pronoun as well as the personal subject pronoun are

more likely to be influenced by information structure than the personal object pronoun. Further, the demonstrative subject pronoun is affected more by information structure than the personal subject pronoun.

This leads us to our next research question addressing whether changes in information structure are processed differently in the different groups under investigation. Particularly, we ask whether children and L2 speakers would benefit in their referent choices when changes in information structure shift their attention to one of the referents.

For the first study, we saw clear differences in processing between children and adults. As mentioned above, we argue that children's discourse processing is local, while adults' is global. These different processing strategies lead to difference in how the information structure effect unfolds. For children, the use of *it*-clefts results in a narrow focus interpretation (i.e., only the clefted referent is in focus), whereas for adults, it signaled broad/dual focus interpretation (i.e., the clefted referent and the succeeding referent are in focus). A narrow focus interpretation means for pronoun processing that the object pronoun referent receives more attention than the reflexive referent. A broad/dual focus interpretation would mean both of the preceding referents are in focus, but since the reflexive referent is mentioned again, attention shifts to that referent. However, ultimately the referent preferences during online processing did not affect offline choices. This does not imply that children's referent choices would not benefit from focus shifting attention to one referent. On the contrary, we argue that the reason why we do not see the effect on the final interpretation is due to the prononminal system being more restricted in German when compared to e.g. Dutch. Because Dutch allows referential links with object pronouns that are not possible in German, the referential function makes it easier for German children to select the correct referent than for Dutch children. This is where information structure comes into play. By marking the intended referent with focus, attention is on that referent which renders that referent more likely to be selected as the object pronoun's referent (cf. van Rij et al., 2016; Spenader et al., 2009). In fact, this is also what we see in the L2 speakers in study 3. Focus marking guides their referent selection. Considering that our L2

speakers in the first experiment were not highly proficient, this could mean that L2 speakers with lower proficiency levels rely more on information structure cues during their interpretation of pronouns than L1 speakers (Patterson et al., 2017, cf.). On the other hand, it would not explain why the preferences only emerge for two conditions (focused subject referent with the personal pronoun and focused object referent with the demonstrative pronoun) unless L2 speakers are aware of the different functionalities of personal and demonstrative pronouns. Another explanation could be that some L2 speakers have a preference to resolve the personal pronoun towards the subject referent and the demonstrative pronoun towards the object referent. These preferences are enhanced when the subject referent aligns with focus marking and when the object referent aligns with focus marking. But when they are disjoined, preferences are reduced. For accented pronouns, L1 and L2 speakers align with their preference for the subject referent with object pronouns. It is not clear whether the referent preference was reversed with accented subject pronouns only for L1 speakers or whether this was the case for L2 speakers as well, but the effect was lost due to L2 speakers' L1. In any case, information structure affected referent selection in a similar way in L1 and L2 speakers for at least object pronouns.

In summary, changes in information structure can have different effects on processing and referent selection in children, L1 and L2 adult groups. Children and L2 speakers may be affected similarly in that their processing/referent selection was can benefit from focus marking of the intended referent to perform more like adult native speakers. That is, for children, focus marking guided attention towards the focused referent even though it did not affect the children in our study, it may aid in referent selection in languages where the referential function of the object pronoun is less restricted. For L2 speakers focus marking in combination with grammatical role resulted in a bias for the focused referent. Moreover, the results of L2 speakers indicate that focus marking does not always lead to a more native-like performance. There was no referent preference for the personal pronoun in the object focus condition, whereas native speakers preferred the subject referent (Hert et al., Chapter 3 in this thesis). Taken together it would suggest that

focus marking guides attention to the focused referent, yet final selection may be still modulated by other factors, such as referential function or grammatical role.

Finally, we want to address the question to which degree information structure affects the processing and interpretation of the different types of pronouns. For online processing, study 1 and study 2 provide evidence that the processing of pronouns is affected when changes in information structure appear. For study 1, processing for both object pronouns and reflexives is influenced by information structure. For children, that meant more looks to the clefted referent than to the non-clefted referent, for adults it resulted in more attention towards the non-clefted referent. However, it is not clear if the effect equally modulated processing of reflexives and object pronouns or whether one was more sensitive than the other. Since the contexts they occurred in were not ambiguous and they cannot be linked to the same referent, we cannot infer the degree the effect ultimately had from offline interpretation. Moreover, we do not know if the effect would have unfolded differently on the reflexive, if the reflexive's referent had been clefted as well, since only the pronoun's intended referent was clefted in our items. For study 2, information structure had a clear effect on online processing, as the focused referent was attended more than the non-focused referent. In the offline data, the personal subject pronoun was influenced by focus marking, but it did not change its subject preference. On the other hand, the demonstrative pronoun changed its preference for the object referent when focus was on the subject resulting in the subject referent being selected more often than with object focus. Therefore, we can conclude that the personal subject pronoun is affected more by the grammatical role than the demonstrative subject pronoun, whereas the demonstrative subject pronoun is more sensitive to focus marking than the personal pronoun (Bader & Portele, 2019a,b; Ellert, 2013; Ellert et al., 2011; Kaiser & Trueswell, 2008). In contrast, for L2 speakers both subject pronouns seemed to be similarly sensitive to focus marking which is in line with previous findings (Ellert et al., 2011; Patterson et al., 2017). Although previous research also suggests that with higher proficiency levels, L2 speaker develop a sensitivity for the different factors influencing the two pronoun types (Ellert et al., 2011; Patterson et al., 2017), we did not see this effect in our data

possibly because of the relatively low levels of proficiency. For the comparison of personal subject and object pronouns in L1 speakers, when unaccented, both pronouns are influenced by grammatical role, specifically the subject referent. When accented, subject and object pronoun are affected differently. The change in information structure resulted in the subject pronoun being linked preferably to the object referent. For unaccented and accented object pronouns, no difference was observed for referent selection. This difference in referent preference for accented pronouns may indicate that object pronouns are influenced by information structure to a lesser degree than subject pronouns. This hypothesis is supported by another finding in the experiment, namely order of mention of the object pronoun not affecting referent choice. Abashidze et al. (2022) argue that moving pronouns to the first-mention position (Vorfeld-position) would give rise to topicality effects. This would imply that there should be a difference in referent selection for object pronouns occurring in first or second position. This idea is not supported by our findings.

To conclude, during online processing all pronouns are sensitive to change in information structure which mostly led to more attention towards the referent in focus. For offline interpretation, pronouns vary in the degree to which they are sensitive to information structure effects which is in line with previous findings for personal and demonstrative subject pronouns (Bader & Portele, 2019a,b; Ellert, 2013; Ellert et al., 2011; Kaiser & Trueswell, 2008). The new finding is that it also applies to object pronouns. Our findings suggest the object pronoun to be as sensitive to information structure cues as the subject pronoun. Moreover, regarding second language acquisition, results from study 3 point out that sensitivity of pronouns is also modulated by L2 speakers' proficiency.

5.3   Limitations and Future Directions

In what follows, we will point out several limitations of our studies and discuss how to address them in future research.

The first study in Chapter 2 examined the effect of information structure on ob-

ject pronouns and reflexives in German children and adults. In the discussion, we point out that our results suggest children's processing differs from that of adults in that children take the local discourse into account, whereas adults rely on the global discourse. Including the results of Spenader et al. (2009) and van Rij et al. (2016), we argue that this is the underlying source of why children's processing of object pronouns differs from adults'. However, the determining factor is the referential functionality of the pronoun itself which can prevent children from incorrectly linking the object pronoun to the preceding subject referent. We further propose that when referential functionality in a language is less restrictive, then children rely more on information structure for the interpretation of the pronoun as this kind of information is readily available in the discourse. To test our theory, future research would need to examine the effect of information structure in a language like Dutch or English.

In the second study in Chapter 3, even though our results of the first experiment suggest that online processing preferences do not necessarily reflect offline referent selection preference (cf. Blything et al., 2021b; Schumacher et al., 2016, 2017), using eye-tracking to analyze online processing of the different subject pronouns in experiment 2 may have given us insight on the processing of the demonstrative pronoun and how it unfolds over time as well as whether it would differ from the processing of personal subject pronouns.

While the second experiment of the third study revealed differences in subject and object pronouns' sensitivity to information structure, future research may also want to include demonstrative object pronouns. This would allow for a more detailed understanding of the different pronouns' sensitivity to different linguistics factors and how it interacts with their referential functions.

Another rather unexpected outcome of the second experiment in the third study is that the results for accented pronouns but also for the unaccented object pronoun do not converge with previous findings (cf. Abashidze et al., 2022; Kehler et al., 2008; Mozuraitis & Heller, 2017; Sauermann & Gagarina, 2017; Taylor et al., 2013). In our explanation we link the different results to differences in the discourse contexts preceding the pronoun. While our experimental items include three referents,

most other studies only use context with two referents (Patterson & Schumacher, 2021, see also) which may affect the information structure setup. Therefore, for future research, we propose to explicitly examine effects of number of referents included in the discourse context.

For the third study in Chapter 4, our findings were restricted by the size and composition of the data set limiting the complexity of our models. Especially for L2 speakers, we were not able to analyze the potential relationship between proficiency and L1 influence and how it may affect their referent selection. This means for future studies to either immensely increase the number of participants or to limit the number of L2 participants' L1s. However, limiting the number of L1s may also lead to over-generalizations of the effect of L1 on L2 pronoun resolution. Lastly, another limitation may lie in our experiments with L2 speakers where we only collected offline selection data. Roberts et al. (2008) not only investigated offline referential choice, but also included an online measure to investigate processing in L2 speakers with different L1s (German and Turkish). Their findings are surprising in that both L2 groups' online processing showed a similar pattern, but their referent choices diverged. Specifically, for the second experiment, including an online measure like eye-tracking may have helped to understand whether referent selection of accented subject pronouns was influenced by L2 speakers' L1. This, however, may have also been restricted by composition and sample size of our data set as mentioned earlier. In any case, while online preference may not always transfer to offline preferences, they could still give insight into what processing patterns underlie a certain referent choice.

5.4   Conclusion

Although pronoun resolution research has a long history, investigations – such as the current ones – are still providing new findings. The current dissertation contributed by, most prominently, adding support to existing research on the different functionalities of different pronouns, but also extending existing research by adding a more detailed picture of the sensitivities of different pronouns as well

as extending the current form-specific account (Kaiser & Trueswell, 2008) to object pronouns. Further, our findings indicate that prominence is not the explanation as to why a certain referent is linked to a pronoun. While focus draws attention to the focus-marked referent – which may be argued to reflect a referent's prominence – it is not the decisive factor for the pronoun's interpretation. That said, for some pronoun types, like the demonstrative subject pronoun, focus, alongside objecthood, plays an important role. In addition, the use of information structure modulations has given rise to the hypothesis that the underlying issue why children can struggle with the interpretation of object pronouns lies in a combination of the object pronoun's referential function and their local processing. Lastly, specifically in the second and third study we carefully controlled prosody for our information structure manipulations which has not been accounted for in previous research.

As for the different factors – more specifically information structure and grammatical role – affecting the various pronoun types under investigation, we can sum up the effects in the following rankings:

(16) Sensitivity towards information structure effects:

demonstrative subject pronoun > personal subject pronoun > personal object pronoun

(17) Sensitivity towards grammatical role effects:

personal subject pronoun = personal object pronoun > demonstrative subject pronoun

In (16), demonstrative subject pronouns are more sensitive to information structure than personal subject pronouns which in turn are more sensitive than personal object pronouns. In (17), personal subject pronouns are as sensitive to grammatical role as personal object pronouns, but more sensitive than demonstrative subject pronouns.

Besides these theoretical contributions, the studies also made a methodological contribution. Namely, we argue for eye-tracking studies to include not only the target segment, but also add the segment preceding and/or following the target segment. We have found by doing so, it enabled us to receive a better under-

standing of the underlying issues. This is not only useful for pronoun studies, but rather, generally, studies investigating processing may benefit from it. Moreover, our studies further suggest that the comprehension of pronouns cannot be simply read off the time course of the gaze data. Both of the eye-tracking experiments have shown that the referent that is attended at the time of the pronouns is not necessarily the referent that is selected for the final interpretation. We therefore propose when using the *Visual World Paradigm* to include a task that taps into the final interpretation.

Bibliography

van den Akker, S., Hoeks, J., Spenader, J. K., & Hendriks, P. (2012). Is the Dutch Delay of Principle B Effect dependent on verb type? *Linguistics in the Netherlands*, *29*, 1–14.

Abashidze, D., Gagarina, N., & Bittner, D. (2022). The processing of the ambiguous object pronoun in German monolingual and highly proficient L2 speakers with L1 Georgian. *International Journal of Bilingualism*, (p. 136700692211134).

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, *19*.

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition*, *6*, 81–96.

Akmajian, A., & Jackendoff, R. (1970). Coreferentiality and Stress. *Linguistic Inquiry*, *1*, 124–126.

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, *106*, 748–765.

Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., & Charles, C. (2002). Null vs. overt pronouns and the Topic-Focus articulation in Spanish. *Italian Journal of Linguistics*, *14*, 151–169.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247–264.

Aravind, A., Hackl, M., & Wexler, K. (2018). Syntactic and Pragmatic Factors in Children's Comprehension of Cleft Constructions. *Language Acquisition*, *25*, 284–314.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics*, *24*, 65–87.

Ariel, M. (1990). *Anaphoric noun-phrase antecedents*. London: Croom Helm.

Ariel, M. (2001). Accessibility Theory: An Overview. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text Representation: Linguistic and Psycholinguistic Aspects* (pp. 29–87). Amsterdam, Netherlands: John Benjamins Publishing Company.

Arnold, J. E. (2001). The Effect of Thematic Roles on Pronoun Use and Frequency of Reference Continuation. *Discourse Processes*, *31*, 137–162.

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Linguistics and Language Compass*, *4*, 187–203.

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes*, 22, 527–565.

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*, 521–536.

Arnold, J. E., Kaiser, E., Kahn, J. M., & Kim, L. K. (2013). Information structure: linguistic, cognitive, and processing approaches. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*, 403–413.

Arnold, J. E., & Lao, S. Y. C. (2015). Effects of psychological attention on pronoun comprehension. *Language, Cognition and Neuroscience*, *30*, 832–852.

Arosio, F., Yatsushiro, K., Forgiarini, M., & Guasti, M. T. (2012). Morphological Information and Memory Resources in Children's Processing of Relative Clauses in German. *Language Learning and Development*, *8*, 340–364.

Avrutin, S., & Wexler, K. (1992). Development of Principle B in Russian: Coindexation at LF and Coreference. *Language Acquisition*, *2*, 259–306.

Baauw, S., & Cuetos, F. (2003). The Interpretation of Pronouns in Spanish Language Acquisition and Breakdown: Evidence for the "Principle B Delay" as a Non-Unitary Phenomenon. *Language Acquisition*, *11*, 219–275.

Baayen, R. H., van Rij, J., de Cat, C., & Wood, S. (2018). Autocorrelated Errors in Experimental Data in the Language Sciences: Some Solutions Offered by Generalized Additive Mixed Models. In D. Speelman, K. Heylen, & D. Geeraerts (Eds.), *Mixed-Effects Regression Models in Linguistics* (pp. 49–69). Springer, Cham.

Bader, M., & Portele, Y. (2019a). Givenness and the Licensing of Object-First Order in German: The Effect of Referential Form. In A. Gattnar, R. Hö rning, M. Stö rzer, & S. Featherston (Eds.), *Procedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory* (pp. 208–228). T ü bingen: University of T ü bingen.

Bader, M., & Portele, Y. (2019b). The interpretation of German personal pronouns and d-pronouns. *Zeitschrift fur Sprachwissenschaft*, *38*, 155–190.

Bahlmann, J., Rodriguez-Fornells, A., Rotte, M., & M ü nte, T. F. (2007). An fMRI study of canonical and noncanonical word order in German. *Human Brain Mapping*, *28*, 940–949.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*, 457–474.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*.

Baumann, S. (2006). *The intonation of givenness: evidence from German.*. Linguistische Arbeiten: 508. Niemeyer.

Beaver, D., Clark, B. Z., Flemming, E., Jaeger, T. F., & Wolters, M. (2007). When semantics meets phonetics: Acoustical studies of second-occurrence focus. *Language*, *83*, 245–276.

Beaver, D. I., Roberts, C., Simons, M., & Tonhauser, J. (2017). Questions Under Discussion: Where Information Structure Meets Projective Content. *Annual Review of Linguistics*, *3*, 265–284.

Belletti, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language and Linguistic Theory*, *25*, 657–689.

Bergmann, C., Paulus, M., & Fikkert, P. (2012). Preschoolers' comprehension of pronouns and reflexives: The impact of the task. *Journal of Child Language, 39*, 777–803.

Birch, S. L., Albrecht, J. E., & Myers, J. L. (2000). Syntactic focusing structures influence discourse processing. *Discourse Processes*, *30*, 285–304.

Blything, L. P., Iraola Azpiroz, M., Allen, S., Hert, R., & Järvikivi, J. (2021a). The influence of prominence cues in 7- to 10-year-olds' pronoun resolution: Disentangling order of mention, grammatical role, and semantic role. *Journal of Child Language*.

Blything, L. P., Järvikivi, J., Toth, A. G., & Arnhold, A. (2021b). The Influence of Focus Marking on Pronoun Resolution in Dialogue Context. *Frontiers in Psychology*, *12*.

Borghini, G., & Hazan, V. (2018). Listening Effort During Sentence Processing Is Increased for Non-native Listeners: A Pupillometry Study. *Frontiers in Neuroscience*, *12*.

Bosch, P., Katz, G., & Umbach, C. (2007). The non-subject bias of German demonstrative pronouns. In M. Schwarz-Friesel, M. Consten, & M. Knees (Eds.), *Anaphors in Text: Cognitive, formal and applied approaches to anaphoric refernce* (pp. 145–164).

Bosch, P., Rozario, T., & Zhao, Y. (2003). Demonstrative Pronouns and Personal Pronouns: German der vs. er. In *Proceedings of the EACL2003 Workshop on The Computational Treatment of Anaphora*. Budapest.

Boswijk, V., & Coler, M. (2020). What is Salience? *Open Linguistics*, *6*, 713–722.

Bouma, G., & Hopp, H. (2006). Effects of word order and grammatical function on pronoun resolution in German. In R. Artstein, & M. Poesio (Eds.), *Ambiguity in Anaphora Workshop Proceedings* August (pp. 5–12).

Bouma, G., & Hopp, H. (2007). Coreference preferences for personal pronouns in German. *ZAS Papers in Linguistics*, *48*, 53–74.

Braun, B. (2006). Phonetics and phonology of thematic contrast in German. *Language and Speech*, *49*, 451–493.

Brennan, S. E. (1995). Centering Attention in Discourse. *Language and Cognitive Processes*, *10*, 137–167.

Büring, D. (1997). *The Meaning of Topic and Focus*. Routledge.

Büring, D. (2003). On D-Trees, Beans, and B-Accents. *Linguistics and Philosophy*, *26*, 511–545.

Büring, D. (2015). A theory of second occurrence focus. *Language, Cognition and Neuroscience*, *30*.

Büring, D., & Križ, M. (2013). It's that, and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites). *Semantics and Pragmatics*, *6*, 1–29.

Callahan, S. M., Shapiro, L. P., & Love, T. (2010). Parallelism effects and verb activation: The sustained reactivation hypothesis. *Journal of Psycholinguistic Research*, *39*, 101–118.

Carminati, M. N. (2002). *The processing of Italian subject pronouns*. Unpublished doctoral thesis University of Massachusetts, Amherst United States – Massachusetts.

Carreiras, M., Gernsbacher, M. A., & Villa, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin & Review*, *2*, 124–129.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and Topic* (pp. 25–55). New York: Academic Press Inc.

Chen, A., Den Os, E., & De Ruiter, J. P. (2007). Pitch accent type matters for online processing of information status: Evidence from natural and synthetic speech. *Linguistic Review*, *24*, 317–344.

Chen, A., & Lai, V. (2011). Comb or coat: The role of intonation in online reference resolution in a second language. *Sound and Sounds. Studies presented to M.E.H. (Bert) Schouten on the occasion of his 65th birthday*, (pp. 57–68).

Chien, Y. C., & Wexler, K. (1990). Children's Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics. *Language Acquisition*, *1*, 225–295.

Chomsky, N. (1969). *Deep structure, surface structure, and semantic interpretation*. Indiana University Linguistics Club.

Clackson, K., Felser, C., & Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language*, *65*, 128–144.

Colonna, S., Schimke, S., & Hemforth, B. (2012). Information structure effects on anaphora resolution in German and French: A crosslinguistic study of pronoun resolution. *Linguistics*, *50*, 991–1013.

Colonna, S., Schimke, S., & Hemforth, B. (2014). Information Structure and Pronoun Resolution in German and French: Evidence from the Visual-World Paradigm. In B. Hemforth, B. Mertins, & C. Fabricius-Hansen (Eds.), *Psycholinguistic Approaches to Meaning and Understanding across Languages. Studies in Theoretical Psycholinguistics* (pp. 175–195). Cham: Springer International Publishing.

Colonna, S., Schimke, S., & Hemforth, B. (2015). Different effects of focus in intra- and inter-sentential pronoun resolution in German. *Language, Cognition and Neuroscience*, *30*, 1306–1325.

Comrie, B. (1997). Pragmatic Binding: Demonstratives as Anaphors in Dutch. *Annual Meeting of the Berkeley Linguistics Society*, *23*, 50.

Conroy, A., Takahashi, E., Lidz, J., & Phillips, C. (2009). Equal Treatment for All Antecedents: How Children Succeed with Principle B. *Linguistic Inquiry*, *40*, 446–486.

Contemori, C., & Marinis, T. (2014). The impact of number mismatch and passives on the real-time processing of relative clauses. *Journal of Child Language*, *41*, 658–689.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*, 51–57.

Cowles, H. W., Walenski, M., & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi*, *26*, 3–18.

Crawley, R. A., & Stevenson, R. J. (1990). Reference in single sentences and in texts. *Journal of Psycholinguistic Research*, *19*, 191–210.

Crawley, R. A., Stevenson, R. J., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, *19*, 245–264.

Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition*, *7*, 49–59.

De la Fuente, I., & Hemforth, B. (2013). Effects of clefting and left-dislocation on subject and object pronoun resolution in Spanish. In J. Cabrelli Amaro, G. Lord, A. de Prada Pérez, & J. E. Aaron (Eds.), *Selected Proceedings of the 16th Hispanic Linguistics Symposium* (pp. 27–45).

Destruel, E., & Velleman, L. (2014). Refining Contrast: Empirical Evidence from the English it -Cleft. In C. Piñón (Ed.), *Empirical Issues in Syntax and Semantics 10* (pp. 197–214).

Diessel, H. (1999). *Demonstratives: form, function, and gram- maticalization.* volume 42 of *Typological studies in language: 42.* Amsterdam: J. Benjamins.

Diessel, H., & Tomasello, M. (2005). A New Look at the Acquisition of Relative Clauses. *Language*, *81*, 882–906.

Dik, S. C., Hoffman, M. E., de Jong, J. R., Sie, I. D., Stroomer, H., & de Vries, L. (1980). On the typology of Focus Phenomena. In *Perspectives on Functional Grammar* (pp. 41–74). De Gruyter.

Drenhaus, H., Zimmermann, M., & Vasishth, S. (2011). Exhaustiveness effects in clefts are not truth-functional. *Journal of Neurolinguistics*, 24, 320–337.

Ellert, M. (2010). *Ambiguous pronoun resolution in L1 and L2 German and Dutch*. Ph.D. thesis Radbout University Nijmegen.

Ellert, M. (2013). Information Structure Affects the Resolution of the Subject Pronouns Er and Der in Spoken German Discourse. *Discours*, (pp. 3–24).

Ellert, M., Roberts, L., & Järvikivi, J. (2011). Verarbeitung und Disambiguierung pronominaler Referenz in der Fremdsprache Deutsch: Eine psycholinguistische Studie. In C. Spiegel, & A. Krafft (Eds.), *Sprachliche Förderung und Weiterbildung-Transdisziplinär* (pp. 51–68).

Fanselow, G. (2015). Syntactic and Prosodic Reflexes of Information Structure in Germanic. In C. Féry, & S. Ishihara (Eds.), *The Oxford Handbook of Information Structure*. Oxford University Press.

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition*, *96*, 263–284.

Féry, C. (1993). *German intonational Patterns*. De Gruyter.

Féry, C. (2017). *Intonation and Prosodic Structure*. Cambridge: Cambridge University Press.

Féry, C., & Ishihara, S. (2009). The phonology of Second Occurrence Focus. *Journal of Linguistics*, *45*.

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, *36*, 680–703.

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language*, *56*, 357–383.

Frascarelli, M., & Ramaglia, F. (2013). (Pseudo)clefts at the Syntax-Prosody-Discourse Interface. In K. Hartmann, & T. Veenstra (Eds.), *Cleft Structures* (pp. 97–138). Philadelphia: John Benjamins Publishing Company.

Frazier, L., Taft, L., Roeper, T., Clifton, C., & Ehrlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition*, *12*, 421–430.

Frederiksen, J. R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes*, *4*, 323–347.

Frey, W. (2005). Zur Syntax der linken Peripherie im Deutschen. In F.-J. D'Avis (Ed.), *Deutsche Syntax: Empirie und Theorie* January 2005 (pp. 147–171). Göteborg: Acta Universitatis Gothoburgensis.

Frey, W. (2006). Contrast and movement to the German prefield. *The Architecture of Focus*, (pp. 235–264).

de la Fuente, I., & Hemforth, B. (2013). Topicalization and focusing effects on subject and object pronoun resolution in Spanish. *Selected Proceedings of the Hispanic Linguistics Symposium 2012*, (pp. 27–45).

Fukumura, K., & van Gompel, R. P. (2015). Effects of order of mention and grammatical role on anaphor resolution. *Journal of Experimental Psychology: Learning Memory and Cognition*, *41*, 501–525.

Gargiulo, C., Tronnier, M., & Bernardini, P. (2019). The role of prosody in overt pronoun resolution in a null subject language and in a non-null subject language: A production study. *Glossa*, *4*.

Gelormini-Lezama, C., & Almor, A. (2011). Repeated names, overt pronouns, and null pronouns in Spanish. *Language and Cognitive Processes*, *26*, 437–454.

Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, *17*, 324–363.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, *27*, 699–717.

Gernsbacher, M. A., Hargreaves, D. J., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language*, *28*, 735–755.

Gillingham, G. (2013). Focusing on unlikely accented nominals: context, alternatives and implied expectations. In T. Snider (Ed.), *Semantics and Linguistic Theory* (pp. 174–193). volume 23.

Givón, T. (1983). *Topic continuity in discourse: a quantitative cross-language study*. Typological studies in language: v. 3. John Benjamins Publishing Company.

Gordon, P. C., & Chan, D. (1995). Pronouns, Passives, and Discourse Coherence. *Journal of Memory and Language*, *34*, 216–231.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*, 311–347.

Grice, M., Baumann, S., & Benzmü ller, S. (2005). German intonation in autosegmental-metric phonology. In S.-A. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* chapter 3. (pp. 55–83). Oxford: Oxford University Press.

Grice, M., Baumann, S., Ritter, S., & R ö hr, C. T. (2017). GToBI. Übungsmaterialien zur deutschen Intonation.

Grimshaw, J., & Rosen, S. (1990). Knowledge and Obedience: The Developmental Status of the Binding Theory. *Linguistic Inquiry*, *21*, 187–222.

Grodzinsky, Y., & Kave, G. (1993). Do Children Really Know Condition A? *Language Acquisition*, *3*.

Grodzinsky, Y., & Reinhart, T. (1993). The Innateness of Binding and Coreference. *Linguistic Inquiry*, *24*, 69–101.

Grodzinsky, Y., Wexler, K., Chien, Y. C., Marakovitz, S., & Solomon, J. (1993). The breakdown of binding relations. *Brain and Language*, *45*, 396–422.

Grosz, B. J., & Sidner, C. L. (1986). Attention, Intention and the Structure of Discourse. *Computational Linguistics*, *12*, 175–204.

Grosz, B. J., Weinstein, S., & Joshi, A. (1995). Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational linguistics - Association for Computational Linguistics (Print)*, *21*, 203–225.

Guasti, M. T. (2002). *Language Acquisition: The Growth of Grammar*. The MIT Press.

Gundel, J. K. (1974). *Role of topic and comment in linguistic theory*. Ph.D. thesis University of Texas at Austin Bloomington, Ind.

Gundel, J. K. (2003). Information Structure and Referential Givenness/Newness: How Much Belongs in the Grammar? *Journal of Cognitive Science*, *4*, 177–199.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, *69*, 274.

Hamann, C. (2002). *From Syntax to Discourse: Pronominal Clitics, Null Subjects and Infinitives in Child Language.* Studies in Theoretical Psycholinguistics: 29. Springer Netherlands.

Hartshorne, J. K., Nappa, R., & Snedeker, J. (2015). Development of the first-mention bias. *Journal of Child Language*, *42*, 423–446.

Hedberg, N. (2013). Multiple focus and cleft sentences. In K. Hartmann, & T. Veenstra (Eds.), *Cleft Structures* (pp. 227–250). Philadelphia: John Benjamins Publishing Company.

Heim, S., & Alter, K. (2006). Prosodic pitch accents in language comprehension and production: ERP data and acoustic analyses. *Acta Neurobiologiae Experimentalis*, *66*, 55–68.

Hendriks, P. (2014). *Asymmetries between Language Production and Comprehension* volume 42 of *Studies in Theoretical Psycholinguistics*. Dordrecht: Springer Netherlands.

Hendriks, P., Banga, A., van Rij, J., Cannizzaro, G., & Hoeks, J. (2011). Adults' online comprehension of object pronouns in discourse. *Production-Comprehension Asymmetries in Child Language*, (pp. 193–216).

Hert, R., Järvikivi, J., & Arnhold, A. (n.d., Chapter 3 in this thesis). Nothing else matters - He likes subject referents.

von Heusinger, K., & Schumacher, P. B. (2019). Discourse prominence: Definition and application. *Journal of Pragmatics*, *154*, 117–127.

Hickmann, M., & Hendriks, H. (1999). Cohesion and anaphora in children's narratives: a comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language*, *26*, 419–452.

Himmelmann, N. P., & Primus, B. (2015). Prominence Beyond Prosody - A First Approximation. In A. De Dominicis (Ed.), *pS-prominenceS: Prominences in Linguistics* December 2013 (pp. 38–58). Viterbo: DISUCOM Press.

Hruska, C., & Alter, K. (2013). Prosody in Dialogues and Single Sentences: How Prosody can influence Speech Perception. In *Information Structure*. Berlin, Boston: de Gruyter.

Ito, K., Bibyk, S. A., Wagner, L., & Speer, S. R. (2014). Interpretation of contrastive pitch accent in six- to eleven-year-old English-speaking children (and adults). *Journal of Child Language*, *41*, 82–108.

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*, 541–573.

Jacobs, J. (1991). Focus ambiguities. *Journal of Semantics*, *8*, 1–36.

Järvikivi, J., van Gompel, R. P., & Hyönä, J. (2017). The Interplay of Implicit Causality, Structural Heuristics, and Anaphor Type in Ambiguous Pronoun Resolution. *Journal of Psycholinguistic Research*, *46*, 525–550.

Järvikivi, J., van Gompel, R. P., Hyönä J., & Bertram, R. (2005). Ambiguous Pronoun Resolution. *Psychological Science*, *16*, 260–264.

Järvikivi, J., Pyykkönen-Klauck, P., Schimke, S., Colonna, S., & Hemforth, B. (2014). Information structure cues for 4-year-olds and adults: Tracking eye movements to visually presented anaphoric referents. *Language, Cognition and Neuroscience*, *29*, 877–892.

Jun, S.-A. (2005). *Prosodic Typology*. Oxford University Press.

Jun, S.-A. (Ed.) (2014). *Prosodic Typology II*. Oxford University Press.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expe´rimentale*, *47*, 310–339.

Kaiser, E. (2005a). Different Forms Have Different Referential Properties: Implications for the Notion of 'Salience' Elsi Kaiser. In A. Branco, T. McEnery, & R. Mitkov (Eds.), *Anaphora processing: Linguistic, cognitive and computational modelling* (pp. 261–282). John Benjamins Publishing Company.

Kaiser, E. (2005b). When salience isn't enough: Pronouns, demonstratives and the quest for an antecedent. In R. Laury (Ed.), *Minimal reference: The use of pronouns in Finnish and Estonian discourse* (pp. 135–162). SKS Finnish Literature Society.

Kaiser, E. (2011a). Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus. *Language and Cognitive Processes*, *26*, 1625–1666.

Kaiser, E. (2011b). On the Relation between Coherence Relations and Anaphoric Demonstratives in German. *Proceedings of the 2010 Annual Conference of the Gesellschaft für Semantik. Sinn und Bedeutung 15,* (pp. 337–351).

Kaiser, E. (2011c). Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives. *Language and Cognitive Processes*, *26*, 1587–1624.

Kaiser, E., & Trueswell, J. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, *94*.

Kaiser, E., & Trueswell, J. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, *23*, 709–748.

Kameyama, M. (1999). Stressed and unstressed pronouns: Complementary preferences. *Focus: Linguistic, cognitive and computational perspectives*, *1*, 306–321.

Keenan, E. L., & Comrie, B. (1977). Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry*, *8*, 63–99.

Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. CSLI Lecture Notes, 104. CSLI Publications, Stanford, CA.

Kehler, A. (2005). Coherence-Driven Constraints on the Placement of Accent. *Semantics and Linguistic Theory*, *15*, 98–115.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*, 1–44.

Kember, H., Choi, J., Yu, J., & Cutler, A. (2021). The Processing of Linguistic Prominence. *Language and Speech*, *64*, 413–436.

Kharitonova, M., Winter, W., & Sheridan, M. A. (2015). As Working Memory Grows: A Developmental Account of Neural Bases of Working Memory Capacity in 5- to 8-Year Old Children and Adults. *Journal of Cognitive Neuroscience*, *27*, 1775–1788.

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*, 860–897.

Kiss, K. E. (1998). Identificational Focus versus Information Focus. *Language*, *74*, 245.

Klin, C. M., Weingartner, K. M., Guzmán, A. E., & Levine, W. H. (2004). Readers' sensitivity to linguistic cues in narratives: How salience influences anaphor resolution. *Memory & Cognition*, *32*, 511–522.

Knoeferle, P., & Crocker, M. W. (2009). Constituent order and semantic parallelism in online comprehension: Eye-tracking evidence from German. *Quarterly Journal of Experimental Psychology*, *62*, 2338–2371.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*, 243–276.

Krifka, M., & Musan, R. (2013). Information structure: Overview and linguistic issues. *The Expression of Information Structure*, (pp. 1–44).

Kristensen, L. B., Wang, L., Petersson, K. M., & Hagoort, P. (2013). The interface between language and attention: Prosodic focus marking recruits a general attention network in spoken language comprehension. *Cerebral Cortex*, *23*, 1836–1848.

Kügler, F., & Calhoun, S. (2020). Prosodic Encoding of Information Structure. In C. Gussenhoven, & A. Chen (Eds.), *The Oxford Handbook of Language Prosody* (pp. 453–467). Oxford University Press.

Ladd, D. R., & Arvaniti, A. (2023). Prosodic Prominence Across Languages. *Annual Review of Linguistics*, *9*.

Lambrecht, K. (1994). *Information Structure and Sentence Form*. Cambridge University Press.

Lambrecht, K. (2001). A framework for the analysis of cleft constructions. *Linguistics*, *39*, 463–516.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*, 1–12.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*, 325–343.

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*.

Leslie, A. M. (1996). A theory of agency. In D. Sperber (Ed.), *Causal Cognition* (pp. 121–141). Oxford University Press.

Lozano, C. (2018). The Development of Anaphora Resolution at the Syntax-Discourse Interface: Pronominal Subjects in Greek Learners of Spanish. *Journal of Psycholinguistic Research*, *47*, 411–430.

Magnuson, J. S. (2019). Fixations in the visual world paradigm: where, when, why? *Journal of Cultural Cognitive Science*, *3*, 113–139.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2009). Pronoun co-referencing errors: Challenges for generativist and usage-based accounts. *Cognitive Linguistics*, *20*, 599–626.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

McKee, C. (1992). A Comparison of Pronouns and Anaphors in Italian and English Acquisition. *Language Acquisition*, *2*.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, *23*, 477–494.

Meng, M., & Bader, M. (2000a). Mode of disambiguation and garden-path strength: An investigation of subject-object ambiguities in German. *Language and Speech*, *43*, 43–74.

Meng, M., & Bader, M. (2000b). Ungrammaticality detection and garden path strength: Evidence for serial parsing. *Language and Cognitive Processes*, *15*, 615–666.

Mennen, I., & De Leeuw, E. (2014). Beyond segments. *Studies in Second Language Acquisition*, *36*, 183–194.

Mozuraitis, M., & Heller, D. (2017). Discourse coherence and the interpretation of accented pronouns. *Dialogue & Discourse*, *8*.

Neeleman, A., & van de Koot, H. (2015). Word Order and Information Structure. In C. Féry, & S. Ishihara (Eds.), *The Oxford Handbook of Information Structure* (pp. 383–401). Oxford University Press volume 1.

Norberg, K. A., & Fraundorf, S. H. (2021). Memory benefits from contrastive focus truly require focus: evidence from clefts and connectives. *Language, Cognition and Neuroscience*, *36*, 1010–1037.

Okuma, T. (2011). Acquisition of topic shift by L1 English speakers of L2 Japanese. *Proceedings of the 11 Generative Approaches to Second Langauge Acquisition Conference (GASLA 2011)*, (pp. 90–100).

Osaka, M., Komori, M., Morishita, M., & Osaka, N. (2012). Neural bases of focusing attention in working memory: An fMRI study based on individual differences. *The Cognitive Neuroscience of Working Memory*, *7*, 130–139.

Padilla, J. A. (1990). *On the Definition of Binding Domains in Spanish* volume 11. Dordrecht: Springer Netherlands.

Papadopoulou, D., Peristeri, E., Plemenou, E., Marinis, T., & Tsimpli, I. (2015). Pronoun ambiguity resolution in Greek: Evidence from monolingual adults and children. *Lingua*, *155*, 98–120.

Patterson, C., Esaulova, Y., & Felser, C. (2017). The impact of focus on pronoun resolution in native and non-native sentence comprehension. *Second Language Research*, *33*, 403–429.

Patterson, C., & Schumacher, P. B. (2021). Interpretation preferences in contexts with three antecedents: Examining the role of prominence in German pronouns. *Applied Psycholinguistics*, *42*, 1427–1461.

Perdomo, M., & Kaan, E. (2021). Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research*, *37*, 349–375.

Philip, W., & Coopmans, P. (1996). The Role of Referenciality in the Acquisition of Pronominal Anaphora. *Proceedings of North East Linguistic Society*, *26*, 241–256.

Poirier, J., Walenski, M., & Shapiro, L. P. (2012). The role of parallelism in the real-time processing of anaphora. *Language and Cognitive Processes*, *27*, 868–886.

Polinsky, M. (2011). Reanalysis in adult heritage language: New evidence in support of attrition. *Studies in Second Language Acquisition*, *33*, 305–328.

Polinsky, M., & Kluender, R. (2007). Linguistic typology and theory construction: Common challenges ahead. *Linguistic Typology*, *11*, 273–283.

Polio, C. (1995). The Use of Zero Pronouns By Nonnative Speakers of Chinese and the Implications for the Acquisition of Nominal Reference. *Studies in Second Language Acquisition*, *17*, 353–377.

Porretta, V., Kyröläinen, A. J., van Rij, J., & Järvikivi, J. (2016). VWPre: Tools for Preprocessing Visual World Data.

Porretta, V., Kyröläinen, A. J., Rij, J. V., Järvikivi, J., van Rij, J., & Järvikivi, J. (2018). Visual world paradigm data: From preprocessing to nonlinear time-course analysis. *Smart Innovation, Systems and Technologies*, *73*, 268–277.

Portele, Y., & Bader, M. (2016). Accessibility and Referential Choice: Personal Pronouns and D-pronouns in Written German. *Discours*.

Prince, E. (1981). Toward a Taxonomy of Given-New. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223–255). New York: Academic Press Inc. volume 14.

Prince, E. F. (1978). A Comparison of Wh-Clefts and it-Clefts in Discourse. *Language*, *54*, 883.

Pyykkönen, P., Matthews, D., & Järvikivi, J. (2010). Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes*, *25*, 115–129.

R Core Team (2019). RStudio: Integrated Development Environment for R.

R Core Team (2020). *RStudio: Integrated Development Environment for R*. R Foundation for Statistical Computing Vienna, Austria.

Rambow, O. (1993). Pragmatic aspects of scrambling and topicalization in German: A Centering approach. *Workshop on Centering Theory in Naturally-Occuring Discourse*, (pp. 20–28).

Reichle, R. (2010). Near-nativelike processing of contrastive focus in L2 French. In *Research in Second Language Processing and Parsing* (pp. 321–344).

Reichle, R., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French : L2 proficiency effects on ERPs. *Studies in Second Language Acquisition*, *36*, 535–564.

Reinhart, T. (1981). Pragmatics and Linguistics: An Analysis of Sentence Topics. *Philosophica*, *27*, 53–94.

Repp, S., & Drenhaus, H. (2015). Intonation influences processing and recall of left-dislocation sentences by indicating topic vs. focus status of dislocated referent. *Language, Cognition and Neuroscience*, *30*, 324–346.

van Rij, J., Hollebrandse, B., & Hendriks, P. (2016). Children's eye gaze reveals their use of discourse context in object pronoun resolution. In *Empirical Perspectives on Anaphora Resolution* (pp. 267–294). De Gruyter.

van Rij, J., van Rijn, H., & Hendriks, P. (2010). Cognitive architectures and language acquisition: A case study in pronoun comprehension. *Journal of Child Language*, *37*, 731–766.

van Rij, J., Vaci, N., Wurm, L. H., & Feldman, L. B. (2020a). Alternative quantitative methods in psycholinguistics: Implications for theory and design. *Word Knowledge and Word Usage*, (pp. 83–126).

van Rij, J., Wieling, M., Baayen, R. H., & Van Rijn, H. (2020b). itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs.

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, *5*, 1–69.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition*, *30*, 333–357.

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*, 912–927.

Rooth, M. (1985). *Association with Focus*. Ph.D. thesis University of Massachussetts.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics: An International Journal of Semantics and Its Interfaces in Grammar*, *1*, 75–116.

Ruigendijk, E., Friedmann, N., Novogrodsky, R., & Balaban, N. (2010). Symmetry in comprehension and production of pronouns: A comparison of German and Hebrew. *Lingua*, *120*.

Ruigendijk, E., & Schumacher, P. B. (2020). Variation in reference assignment processes: psycholinguistic evidence from Germanic languages. *Journal of Comparative Germanic Linguistics*, *23*, 39–76.

Sæbø, K. J. (2019). Presupposition and Contrast: German 'aber' as a Topic Particle. *Proceedings of Sinn Und Bedeutung*, (pp. 257–271).

Sauermann, A., & Gagarina, N. (2017). Grammatical Role Parallelism Influences Ambiguous Pronoun Resolution in German. *Frontiers in Physiology*, *8*, 1–8.

Sauermann, A., Höhle, B., Chen, A., & Järvikivi, J. (2011). Intonational marking of focus in different word orders in German children. *Proceeding of the 28th West Coast Conference on Formal Linguistics*, (pp. 313–322).

Schimke, S., & Colonna, S. (2016). Native and Nonnative Interpretation of Pronominal Forms: Evidence from French and Turkish. *Studies in Second Language Acquisition*, *38*, 131–162.

Schriefers, H., Friederici, A. D., & Kühn, K. (1995). The Processing of Locally Ambiguous Relative Clauses in German.

Schumacher, P. B. (2005). *The Syntax–Discourse Interface* volume 80 of *Linguistik Aktuell/Linguistics Today*. Amsterdam: John Benjamins Publishing Company.

Schumacher, P. B., Dangl, M., & Uzun, E. (2016). Thematic role as prominence cue during pronoun resolution in German. *Empirical Perspectives on Anaphora Resolution*, (pp. 213–240).

Schumacher, P. B., Piñango, M. M., Ruigendijk, E., & Avrutin, S. (2010). Reference assignment in Dutch: Evidence for the syntax-discourse divide. *Lingua*, *120*, 1738–1763.

Schumacher, P. B., Roberts, L., & Järvikivi, J. (2017). Agentivity drives real-time pronoun resolution: Evidence from German er and der. *Lingua*, *185*, 25–41.

Schwarzschild, R. (1999). Givenness, avoidf and other constraints on the placement of accent. *Natural Language Semantics*, *7*, 141–177.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109–147.

Sekerina, I. A., Stromswold, K., & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language*, *31*, 123–152.

Shoji, S., Dubinsky, S., & Almor, A. (2017). The Repeated Name Penalty, the Overt Pronoun Penalty, and Topic in Japanese. *Journal of Psycholinguistic Research*, *46*.

Sigurjónsdóttir, S., & Hyams, N. (1992). Reflexivization and Logophoricity: Evidence From the Acquisition of Icelandic. *Language Acquisition*, *2*, 359–413.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, *23*, 197–229.

Song, H. J., & Fisher, C. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language*, *52*, 29–57.

Song, H. J., & Fisher, C. (2007). Discourse prominence effects on 2.5-year-old children's interpretation of pronouns. *Lingua*, *117*, 1959–1987.

Spenader, J., Smits, E. J., & Hendriks, P. (2009). Coherent discourse solves the pronoun interpretation problem. *Journal of Child Language*, *36*, 23–52.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic Roles, Focus and the Representation of Events. *Language and Cognitive Processes*, *9*, 519–548.

Stevenson, R. J., Nelson, A. W., & Stenning, K. (1995). The Role of Parallelism in Strategies of Pronoun Comprehension.

Takahashi, C., Kao, S., Baek, H., Yeung, A. H., Hwang, J., & Broselow, E. (2018). Native and non-native speaker processing and production of contrastive focus prosody. *Proceedings of the Linguistic Society of America*, *3*, 35.

Taylor, R. C., Stowe, L. A., Redeker, G., & Hoeks, J. C. (2013). Comprehension of marked pronouns in Spanish and English: Object anaphors cross-linguistically. *Quarterly Journal of Experimental Psychology*, *66*, 2039–2059.

Thompson, L. A., Driscoll, D., & Markson, L. (1998). Memory for visual-spoken language in children and adults. *Journal of Nonverbal Behavior*, *22*, 167–187.

Thornton, R., & Wexler, K. (1999). *Principle B, VP Ellipsis, and Interpretation in Child Grammar*. Number Vol. 31 in Current Studies in Linguistics. Cambridge, Mass: The MIT Press.

Thul, R., Conklin, K., & Barr, D. J. (2021). Using GAMMs to model trial-by-trial fluctuations in experimental data: More risks but hardly any benefit. *Journal of Memory and Language*, *120*, 104247.

Tomaszewicz-Özakın, B., & Schumacher, P. B. (2022). Anaphoric Pronouns and the Computation of Prominence Profiles. *Journal of Psycholinguistic Research*, *51*, 627–653.

Tönnis, S. (2022). It is not the Obvious Question that a Cleft Addresses. In *Language, Logic, and Computation. TbiLLC 2019* (pp. 128–147). Springer, Cham.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*, 89–134.

Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism*, *8*, 257–277.

Umbach, C. (2005). Contrast and information structure: A focus-based analysis of but. *Linguistics*, *43*, 207–232.

Varlokosta, S., Karafoti, P., & Karzi, V. (2000). On the Acquisition of Pronominal Reference in Child−Greek. In *New Directions In Language Development And Disorders* (pp. 97–103). Boston, MA: Springer US.

Venditti, J. J., Stone, M., Nanda, P., & Tepper, P. (2002). Discourse Constraints on the Interpretation of Nuclear-accented Pronouns. *Speech Prosody 2002. Proceedings of the 1st International Conference on Speech Prosody*.

Vilkuna, M. (1995). Discourse Configurationality in Finnish. In K. E. Kiss (Ed.), *Discourse Configurational Languages* Oxford Studies in Comparative Syntax (pp. 244–268). New York: Oxford University Press.

Vogelzang, M., Guasti, M. T., van Rijn, H., & Hendriks, P. (2021a). How Children Process Reduced Forms: A Computational Cognitive Modeling Approach to Pronoun Processing in Discourse. *Cognitive Science*, *45*.

Vogelzang, M., Hendriks, P., & van Rijn, H. (2016). Pupillary responses reflect ambiguity resolution in pronoun processing. *Language, Cognition and Neuroscience*, *31*, 876–885.

Vogelzang, M., Hert, R., & Ruigendijk, E. (2021b). Remnants of the Delay of Principle B Effect in Adults: A New Approach to an Old Problem. In V. Torrens (Ed.), *Syntax Processing* (pp. 228–247). Newcastle: Cambridge Scholars Publishing.

Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, *25*, 905–945.

Wang, B., Wang, L., & Qadir, T. (2011). Prosodic realization of focus in six languages/dialects in China. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 2011)*, (pp. 144–147).

Weskott, T. (2003). *Information Structure as a Processing Guide: The Left Periphery of German Verb-Second Sentences and Its Interpretation in Context*. Ph.D. thesis Leipzig University.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, *70*, 86–116.

Wilson, F. (2009). *Processing at the syntax-discourse interface in second language acquisition*. Ph.D. thesis University of Edinburgh.

Winterstein, G. (2012). What but-sentences argue for: An argumentative analysis of but. *Lingua*, *122*, 1864–1885.

Wood, S. N. (2017). Generalized Additive Models. In *Generalized Additive Models*. Boca Raton: Chapman and Hall/CRC. (2nd ed.).

Xu, X. (2015). The influence of information status on pronoun resolution in Mandarin Chinese: evidence from ERPs. *Frontiers in Psychology*, *6*, 1–17.

Zerbian, S., & Böttcher, M. (2019). Stressed Pronouns in Mono-and Bilingual German. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *19th International Congress of Phonetic Sciences* (pp. 2640–2644). Canberra, Australia: Australasian Speech Science and Technology Association Inc.