Eberhard Karls Universität Tübingen

Seminar für Sprachwissenschaft

Acceptability Judgements
On Contrastive Dialogues Involving Ellipsis:
A Pilot Study

Thesis submitted for the degree of Master of Arts

Supervised by

Prof. Dr. Michael Franke and

Jun.-Prof. Dr. James Griffiths

Summer Term 2023

Miriam Schiele

miriam.schiele@student.uni-tuebingen.de

Matriculation Number: 6098290

Käthe-Kollwitz-Straße 15

72074 Tübingen

Date of Submission: September 20th, 2023

Table of contents

Table of figures

Abstract

This pioneering study explores the factors that influence how native speakers perceive fragments in dialogues involving contrastive focus in German. The auditory and written dialogues differentiated in whether the fragmentary answer included lexical words, namely nouns denoting human referents, or functional words, specifically prepositions. They also varied in whether contrasting words were emphasised or not. A total of 100 participants were asked to rate the fragmentary answers in the dialogues using a 7-point Likert scale to assess their naturalness. This study hypothesised that (1) auditory dialogues will be perceived as more natural compared to their written counterparts, (2) dialogues featuring either prosodically or orthographically emphasised contrasting words will receive higher ratings than those without emphasis, and (3) dialogues incorporating lexical fragments will receive higher naturalness ratings than those containing functional fragments. The results align with the hypotheses, as auditorily presented dialogues bearing pitch accent on the contrasting words receive the highest ratings. However, in contrast to the final hypothesis, dialogues incorporating prepositions as functional fragments were perceived as more natural than those containing lexical contrasting words. These findings shed light on the complex interplay of acceptability ratings, ellipsis processing intricacies, and the identification of correlate-remnant pairings. In this manner, the study not only lays the groundwork for future research using acceptability judgement tasks but also enhances the understanding of language comprehension processes.

1. Introduction

In this chapter, the background and motivation for the present study is provided, outlining the research questions and objectives that guide the investigation. The significance of this study is emphasised, while acknowledging the scope and limitations inherent in the research design.

1.1 Background and motivation

The pronunciation of a sentence holds significance, encompassing not only the stressed words but also the contents emphasised in the conversation. However, it is reasonable to assume that the perception of naturalness by native speakers may not solely depend on the stressed word but also on its associated meaning as well as the structure of the sentence. This becomes particularly intriguing in dialogues that involve contrasts and incomplete sentences as answers, as illustrated in the following example in German (1).

(1)   A: Peter hat AB 18 Uhr im Kino gearbeitet.
               'Peter worked at the cinema FROM 6pm.'
        B: Nein, BIS 18 Uhr.
               'No, UNTIL 6pm.'

(own example)

Speaker B's response in (1) lacks a complete sentence structure. Therefore, to understand the intended meaning of speaker B's response, the reader must construct a complete sentence using both speaker A's preceding utterance and speaker B's response. Hence, the reader can derive the following complete sentence: *Peter worked at the cinema until 6pm*.

However, for the reader to grasp its intended message, they need to first identify the word *bis* 'until' as contrasting with *ab* 'from' and then discern the intended meaning behind speaker B's response. If the reader is not primed for the contrast through orthographic marking, understanding speaker B's response could become even more challenging. This becomes evident in dialogues that do not include any orthographic marking, as demonstrated in example (2).

(2)   A: Peter hat ab 18 Uhr im Kino gearbeitet.

> ‘Peter worked at the cinema from 6pm.’
> B: Nein, bis 18 Uhr.
>   ‘No, until 6pm.’

<div align="right">(own example)</div>

In the dialogue (2), the contrasting words lack emphasis and therefore, the reader is confronted with the contrast in speaker B's answer unexpectedly. It is yet to determine to what extent native speakers encounter difficulties in interpreting dialogues lacking emphasis on the contrastive words such as (2) in comparison to dialogues such as (1) and what other factors are at play during the comprehension of such dialogues.

Therefore, the study examines the impact of modality and fragment type on the acceptability ratings of dialogue involving contrastive focus and fragmentary answers. That is, a comparative analysis is conducted between the dialogues in (1) and (2) and their verbal equivalents. Additionally, while the contrasting words in (1) and (2) are prepositions, possessing functional meaning, the study also explores dialogues containing lexical words, namely nouns denoting human referents, as contrasting elements.

The present paper is subdivided as follows. Chapter 2 delves into the theoretical background of fragments, focus, ellipsis comprehension, and acceptability judgement tasks, providing a more detailed explanation of the hypotheses. Chapter 3 centres on the study design, stimuli, data collection, and participant information. Chapter 4 presents the findings of the study and revisits the hypotheses, while chapter 5 explores and addresses any confounding factors related to the findings and compares the study's findings to previous research. Last, chapter 6 concludes with a summary of the study, discusses its contributions to the field as well as its limitations, and offers insights into potential avenues for future research.

1.2 Research questions and objectives

The present paper aims to determine the most effective medium for reliably obtaining judgements about such dialogues involving contrastive focus and fragmentary answers, pathing the way for future research using acceptability judgement tasks. Through the analysis of various stimulus characteristics, the aim is to enhance the comprehension of how modality, emphasis,

and fragment type collectively influence the perceived acceptability of fragments. This section will present the factors and hypotheses investigated in the present paper.

First, as illustrated by the examples (1) and (2), dialogues incorporating orthographically marked contrasting words are compared to those lacking emphasis. The present paper aims to explore the perceived naturalness of fragmentary responses in these dialogues among native speakers. The first hypothesis posits that dialogues emphasising the contrasting words will be deemed more natural by native speakers. This prediction is grounded in the assumption that emphasising the contrasting elements enhances their prominence, thereby aiding comprehension and resulting in increased acceptability. For an overview of the role of emphasis in sentence comprehension, see chapter 2.3.

Next, a contrast is drawn between (1) and (2) on the one hand and their verbal equivalents one the other hand. That is, the sentences (1) and (2) are recited by native speakers and diverge based on whether they exhibit default intonation or whether they emphasise the contrasting words through prosody. Presenting auditory dialogues is expected to establish an authentic and true-to-life context for fragmentary answers, consequently fostering higher acceptability ratings compared to written stimuli. Thus, the second hypothesis holds that in general, auditory stimuli will be more likely to find acceptance among native speakers than written stimuli.

Last, dialogues such as (1) and (2) are compared to dialogues where the contrastive words carry lexical meaning rather than functional meaning. This entails a shift from using prepositions such as *bis* 'until' and *ab* 'from' to emphasising nouns such as *Bruder* 'brother' and *Vater* 'father', as shown in example (3). A corresponding example (4) is provided as a parallel to (3), albeit without the inclusion of emphasis.

(3) A: Peter hat seinem BRUDER ein Buch geschenkt.
        'Pete gave a book to his BROTHER.'
     B: Nein, seinem VATER.
        'No, his FATHER.'

(own example)

(4) A: Peter hat seinem Bruder ein Buch geschenkt.
        'Pete gave a book to his brother.'

B: Nein, seinem Vater.
'No, his father.'

<div align="right">(own example)</div>

As will be outlined in chapter 2.3, lexical words are commonly assumed to be more prominent in dialogues and hold greater significance in understanding the meaning of an utterance. Thus, the third hypothesis posits that stimuli with lexical words in contrastive focus would receive higher acceptability ratings than stimuli with functional words in contrastive focus.

1.3 Significance of the study

In the following section, the significance of the present study is demonstrated by exploring the implications of the findings for theoretical frameworks and practical applications. The present study examines several factors that have received limited or no comprehensive investigation in previous research.

First, while the significance of orthographic marking has been explored in other linguistic contexts, its role in the comprehension of contrastive focus remains unexplored. Chapter 2.3 provides an overview of the prior investigations conducted on the topic of orthographic marking in said other linguistic contexts.

Second, the research gap concerning the significance of emphasis in comprehending fragmentary answers necessitates further research. While numerous studies have delved into the realms of prosody and ellipsis comprehension, limited information is available concerning the impact of pitch accent placement on the perceived naturalness of such fragmentary responses. In essence, it remains uncertain whether the placement of pitch accent on the correlate of the fragmentary answer is essential, or if the naturally occurring default intonation of the preceding utterance suffices for comprehending the fragmentary answer.

Additionally, the majority of the existing studies, as outlined in chapter 2.3, have predominantly focused on the English language, disregarding the possibility of crosslinguistic distinctions in the positioning of sentence accent and comprehension of contrastive fragments. Therefore, the present study

adds to the existing body of knowledge, allowing for crosslinguistic comparisons.

Next, earlier investigations have exclusively centred on lexical contrastive responses, primarily in the form of proper names, as illustrated in chapter 2.2 and 2.3. These studies have formed the foundation for theories regarding the processing of such structures. However, it is plausible that disparities exist between processing proper names and processing other words that denote inanimate referents or possess functional meaning. Consequently, the processing theories must account for such differences.

Last, the ongoing debate about formal and informal methods of conducting acceptability judgements, as outlined in chapter 2.4, underscores the necessity for obtaining more dependable data adhering to scientific standards. The present study aims to fulfil this requirement by presenting formally acquired data.

Therefore, this study represents a pioneering endeavour, systematically investigating and contrasting different media for gathering acceptability judgements pertaining to contrastive fragments. By delving into the distinctions between stimuli with varying levels of emphasis and differing fragment types, this research will not only enhance the comprehension of fragmentary constructs but also offer valuable insights for forthcoming studies in the fields of linguistics and psycholinguistics.

1.4 Scope and limitations

The following subchapter addresses the limitations inherent in the research design, acknowledging the potential constraints and scope of the investigation.

First, the investigation focuses exclusively on clausal ellipsis, disregarding semantic and pragmatic ellipsis. This decision was made to ensure that the analysis remains well-defined and manageable within the given scope of the study. Semantic and pragmatic ellipsis could be potential avenues for future research, but they fall outside the boundaries of the present investigation.

Second, the study is limited to exploring syntactic ellipsis solely within the clausal context. While ellipsis can occur at various linguistic levels and

analyses of these phenomena offer valuable insights, they lie beyond the current investigation's scope. Therefore, for the sake of depth and coherence, it was opted to concentrate on clausal ellipsis only. Future studies could explore other types of ellipsis to gain a more comprehensive understanding of ellipsis phenomena.

Additionally, the study does not extend its analysis to dialectal variations, differences between age groups, genders, or any other participant-specific features. Previous research indicated that except for age and geography, such sociolinguistic features do not significantly influence the acceptability ratings (cf. Delbar, 2019).

Within its defined scope, the research design utilises acceptability judgement tasks. However, this method inherently presents limitations. Participant responses may be influenced by individual linguistic competence, biases, and subjective interpretations, introducing potential sources of uncertainty. Although measures to mitigate these issues were taken, such as ensuring a diverse participant pool, providing example dialogues in the introduction to the study, and utilising statistical analysis, it is essential to recognise these inherent limitations.

In conclusion, the present study is bounded by specific limitations and a carefully defined scope. By recognising these limitations, the present study ensures the reliability of its findings and identifies potential avenues for future research.

2. Literature review

In this chapter, the fragment theory is explained, covering its linguistic foundations. Next, the terminology related to focus is explained, encompassing focus in general, contrastive focus and contrastive fragments. Furthermore, the role of emphasis in the comprehension of fragmentary answers is discussed and research on written and auditory stimuli as well as functional and lexical stimuli is reviewed. Last, the methodological approach of acceptability judgement tasks is discussed.

2.1 Fragment theory and its linguistic foundations

In the following section, the concepts of ellipsis and fragment, which can be seen as a more specific form of ellipsis, are explained. Additionally, an example for a fragment is examined, and the licencing conditions for fragments are outlined briefly.

In a general sense, ellipsis is used to refer to an interface phenomenon that occurs among syntax, semantics, and information structure, where linguistic material is omitted (cf. Lobeck, 1995; Winkler, 2019). In other words, "there is meaning without form" (Merchant, 2019, p. 19). While there are numerous types of ellipsis, an in-depth analysis of each of those types would go beyond the scope of this paper. Instead, the present paper focuses on a specific type of ellipsis, i.e. fragments. For a first understanding of fragments, consider (5).

(5) Abby and Ben are at a party. Abby asks Ben about who their mutual friend Beth is bringing as a date by uttering: "Who is Beth bringing?" Ben answers:
"Alex."

(Merchant, 2005, p. 661)

Ben's answer *Alex* in (5) only consists of one word and yet, the reader can easily derive that Ben's utterance is intended to convey that Beth is bringing Alex. Such short answers are called fragments (cf. Merchant, 2004).

In the following, Merchant's (2004) theory is adopted, according to which fragments move to the clause-peripheral position and the concomitant ellipsis operation deletes the remaining constituents. Therefore, the remnant of ellipsis is the part of the underlying clause that survives ellipsis (cf. Griffiths et al., 2023). Hence, Ben's answer *Alex* is the only pronounced constituent of its underlying clause *Beth is bringing Alex*. For an in-depth explanation of how fragmentary answers are comprehended, see chapter 2.3.

Similar to other elliptical structures, fragments must meet specific licencing conditions. For instance, the omission of material is only permissible if the material is given within the discourse (cf. Winkler, 2019, Merchant, 2001). That is, if a part of a sentence is anaphoric to linguistic material in

the previous discourse or is inferred from it, that part is either deaccented or not phonetically realised (cf. Winkler, 2019).

In the example (5), the given material *Beth is bringing* from the clause *Beth is bringing Alex* is eligible for omission since it can be deducted from the prior conversation. However, who Beth is bringing is not clear from the context. Therefore, the information *Alex* is new to the context and consequently, cannot be diminished.

2.2 Contrastive focus

Before examining contrastive fragments, the notion of *focus*, particularly, *contrastive focus* is explained and illustrated with examples. Subsequently, the chapter directs its attention specifically towards fragments that incorporate contrastive focus.

The notion of *focus* is commonly associated with the element of an utterance that adds new information to the discourse (cf. Lambrecht, 1994). While more elaborate definition exceeds the scope of this investigation, it is not necessary to grasp the concept of contrastive focus, which can be defined as follows.

> (6) Contrastive focus represents a subset of contextually or situationally "given" alternative elements for which the predicate phrase can potentially hold, and spells out this subset as the one for which the predicate actually hold.
> (Griffiths & Lipták, 2014, p. 200, quotation marks in original)

However, the alternative denotations must be of the same type and mutually exclusive (cf. Krifka, 2008; Wagner, 2012). An example of contrastive focus is illustrated in (7), where the brackets subscripted with *F* indicate the constituent in focus.

> (7) A: Mary stole the cookie.
>   B: No, [Peter]$_F$ stole the cookie!
> (adapted from Krifka, 2008, p. 252)

The sentence in (7A) is called an antecedent clause, while speaker B's answer is named a contrastive utterance, since *Mary* and *Peter* are contrasting. Moreover, contrastive focus is placed on *Peter* in accordance with the definition in (6).

As can be seen in (7B), the contrastive focus placed on *Peter* serves to emphasise Peter as an "alternative answer to an explicit or implicit statement provided by the previous discourse/situation" (Wagner, 1999, p. 1529). Placing focus on *Peter* serves to convey implicitly that the assertion made by speaker A is erroneous, indicating that Mary did not stole the cookie, but rather, it was Peter who did. Therefore, speaker B's utterance functions as a correction to speaker A's statement.

After examining (7B), which represents a complete sentence as correction, now consider its fragmentary equivalent (8B).

(8)   A: Mary stole the cookie.
       B: No, [Peter]$_F$!

(adapted from Krifka, 2008, p. 252)

The full correction in (7B) and the elliptical correction in (8B) have the same semantic meaning and pragmatic function, despite of the fact that (8B) only consists of a fragment. This is because fragmentary answers are assumed to be structurally identical to full sentences, as explained in chapter 2.2.

Contrastive fragments represent a distinct subset within the category of fragments. Unlike other fragments, they include "an explicit relation of contrast between the elliptical remnant and its correlate in the antecedent clause" (Griffiths & Lipták, 2014, p. 199, emphasis omitted). The dialogue in (8) exemplifies this phenomenon, as the contrast between *Peter* and *Mary* denote opposing referents. Therefore, contrastive fragments can be used for corrections and invariably involve contrastive focus (cf. Griffiths & Lipták, 2014; Krifka, 2008).

Moreover, contrastive focus has often been characterised phonologically with a L+H* pitch for German (cf. Wagner, 1999). This pattern is defined in the Contrastive Remnant Condition (henceforth, CRC), as formulated in (9).

(9) Given information licenses a contrastive focus interpretation of the remnant(s). The contrastive remnant(s) must be assigned a strong contrastive pitch accent.

(Winkler, 2019, p. 363)

Adhering to the terminology used in the ToBI-framework (cf. Beckman & Ayers, 1997), the bitonal L+H* pitch accent is a combination of a low tone (L) and a high pitch (H*). When combined as L+H* pitch accent, it indicates a pitch accent pattern where a low tone is followed by a high tone with a rising pitch. This pattern is often associated with a pitch accent that starts low and rises to a high pitch, creating a pattern of emphasis and prominence on the syllable. It is used to mark significant or accented syllables within an utterance (ibid.).

The prosodic markings on the remnant and correlate indicate their contrast with each other (cf. Winkler, 2019; Rasekhi & Harris, 2021). Adhering to the CMC formulated in (5), contrastive focus is placed on *Peter* in both (7B) and (8B). The importance of not only pitch accents on elements in contrastive focus but the role of emphasis in the comprehension of fragmentary answers in general is explained in the following subchapter.

2.3 The comprehension of fragmentary answers

This section explores the processes behind the comprehension of fragmentary answers involving contrastive focus. First, the process of the comprehension of elliptical utterances is explained. Then, the extent to which emphasis and semantics play a role in said process is discussed.

Comprehending any utterance requires the evaluation and processing of information, i.e. organising lexical meaning and understanding syntactic structures (cf. Harris & Carlson, 2018). This is particularly difficult for elliptical utterances such as dialogues involving fragmentary answers, since the syntactic structure is not present and has to be constructed by the processor, i.e. hearer or reader of the dialogue (cf. Phillips & Parker, 2014). For comprehending elliptical utterances, the processor must finish the following three basic tasks, described in (10).

(10)   Basic tasks of the processor in ellipsis processing:

1. Parse the remnant by constructing the appropriate phrase structure for the remnant given the input.
2. Locate the correlate, if any, from the antecedent clause.
3. Construct the elided phrase by regenerating or copying a structure at Logical Form.

(Harris & Carlson, 2018, p. 485)

In the example (8B), that means that first, *Peter* is identified as remnant. Next, *Mary* is retrieved as correlate to *Peter*, as it appears to be a suitable contrasting denotation. Last, the elided phrase is contrasted, i.e. *Peter stole the cookie*.

Even though the significance of information structure, specifically the differentiation between previously mentioned and newly introduced information, has been demonstrated in chapter 2.1, it remains essential to explore additional factors that influence the interpretation of contrastive fragments. Hence, the subsequent section examines the importance of parallelism, the correlate's position, and the role of orthographic and prosodic emphasis.

Morphological Parallelism refers to the observation that "the processor favors correlate-remnant pairings for which the DPs are maximally similar along semantic and morphological dimensions" (Rasekhi & Harris, 2021, p. 7), where DP stands for *determiner phrase*. As a method for resolving sentence ambiguity and determining the appropriate correlate for the remnant in elliptical structures, it stands as a valuable tool, especially in the context of German elliptical structures. This is due to the case marking present on all DPs in the German language. Consequently, the processor can readily pinpoint the accurate correlate-remnant pairing by identifying matching case markers (cf. Rasekhi & Harris, 2021).

However, parallelism between the correlate and remnant is not limited to their morphological properties. Frazier et al. (1984) found that if the antecedent clause and the elliptical structure both had similar syntactic structures and thematic roles, elliptical structures can be processed faster than dissimilar sentences.

Another factor that influences how elliptical structures are interpreted is the locality of the correlate. That is, the processor tends to choose the

nearest constituent in the preceding clause to be contrasted with the remnant (cf. Harris, 2015). For instance, consider (11).

(11)   John didn't take the poodle to the park, let alone…
        (a) the zoo. (Local contrast with *the park*)
        (b) the pug. (Nonlocal contrast with *the poodle*)
                        (Harris & Carlson, 2016, p. 6, parentheses in original)

Therefore, a contrast between correlate and remnant such as in (11a) is preferred because the correlate and remnant are structurally closer together than in (11b). If this locality bias is violated, the processing of the sentence is slowed down. That is, the eye-tracking study by Frazier and Clifton (1998) indicated that ambiguous sentences with two potential correlates were read faster than unambiguous sentences. Furthermore, a forced-choice study showed that the local DP of an ambiguous sentence was selected as appropriate correlate more often than the nonlocal DP. In addition, Harris and Carlson (2016) conducted self-paced reading studies to explore how elliptical structures involving *let alone*-phrases are processed. Their results indicate that sentences that do not adhere to the locality bias incur a processing cost (ibid.).

Moreover, Rasekhi and Harris (2021) investigate to what extent those factors facilitate the comprehension of elliptical clauses. Their study indicates that in Persian, whether the correlate and remnant have maximally similar semantic and morphological properties seems to play a greater role than the position of the remnant when finding the appropriate contrasting correlate for the remnant. In addition, a processing cost might also occur if the DPs differ in number (cf. Black et al., 1985).

An alternative reasoning for the locality bias is that the closest DP is selected as the correlate not due to its structural proximity, but because object DPs naturally carry the primary sentence accent in SVO sentences, even during silent reading (cf. Rasekhi & Harris 2021). This holds true for German as well, as the same nuclear accent is also placed on the object DPs in present perfect sentences in the German language (cf. Féry, 2011; Féry & Herbst, 2004). Hence, the significance of orthographic and prosodic

emphasis becomes another pivotal element influencing the perception of contrastive fragments.

Previous research has studied the presentation of words such as capitalisation or colour highlighting can impact comprehension. By employing different forms of orthographic marking, these studies have investigated different fields of linguistics such as lexical access (cf. Opitz & Bordag, 2022), and second language acquisition (cf. Meurers et al., 2010).

Opitz and Bordag (2022) investigated whether initial letter capitalisation of German nouns leads to faster processing of word-class information, as capitalisation distinguishes nouns from other word classes in German. Their findings indicate that as hypothesised, such orthographic cues facilitate word-class specific lexical access. Nevertheless, the utilisation of initial letter capitalisation in the study conducted by Opitz and Bordag (2022) differs from the capitalisation approach employed in the current study, where the entire word is rendered in uppercase letters. Additionally, it is important to note that the research inquiries are situated within distinct linguistic domains.

Introducing a tool designed to adjust web content to support language learners through methods such as colour highlighting, Meurers et al. (2010) explore the effectiveness of this tool in assisting learners' comprehension and vocabulary acquisition. According to Meurers et al. (2010) the orthographically modified web pages positively impact learners' language skills.

Therefore, it can be assumed that the orthographic modifications applied to contrasting words in dialogues in the present study could potentially enhance readers' grasp of the contrastive response. It is crucial, however, to recognise that Meurers et al. (2010) focus on second language acquisition, while the current study is concerned with the comprehension of contrastive focus.

Furthermore, Franck et al. (2003) examine the impact of orthographic factors on grammatical agreement, exploring the interplay between orthography and modality. Specifically, they contrast singular and plural forms that share the same pronunciation but have different spellings, with forms that share both the same written and spoken representations. Notably, their findings reveal that the influence of orthography is observable solely in written

stimuli and not in their spoken equivalents, suggesting that the form effects of orthography are modality-specific and influence grammatical encoding differently depending on whether the language is being written or spoken.

While the study by Franck et al. (2003) offers an intriguing correlation between orthography and marking, which is also pertinent to the present study's interests, it is worth noting that their research only encompassed the orthographic representation of words. It did not involve any additional markers such as capitalisation or colours. Consequently, their study provides only limited insights into the research questions addressed in the current study.

It is evident that although research on orthography exists in other linguistic domains, its influence on the processing of contrastive fragments remains to be established. Moreover, the present study not only examines orthographic cues but also explores prosodic markers. Consequently, the remaining part of the subchapter is concerned with the research concerning the impact of prosody on comprehension.

Prosody is known for impacting language processing (cf. Warren 1999). However, the influence of intonation and emphasis is especially intriguing for the processing of ellipsis sentences and for structures involving contrastive focus. Carlson et al. (2009) investigated whether pitch accent affects how ambiguous replacive sentences such as (12) are interpreted.

(12)　a.　ROGER insisted that Alice was reliable // not ANDREW[.]
　　　 b.　Roger insisted that ALICE was reliable // not ANDREW[.]
　　　 c.　ROGER insisted that ALICE was reliable // not ANDREW.
(Carlson et al., 2009, p. 1077)

Indeed, they found that pitch accent significantly influenced participants' choice of the correlate of the replacive, i.e. whether the expression *Andrew* is used to replace *Roger* or *Alice*. There are two main differences between the sentences used in the study by Carlson et al. (2009) and the sentences used in the present study. First, the former uses replacive sentences, while the latter uses fragmentary answers. Since both are forms of ellipsis, it is still reasonable to hypothesise significant differences comparing dialogues with and without emphasis on the contrasting words. Second, the former uses ambiguous sentences, while the latter uses sentences disambiguated

using case marking. Although the prosodic marking is not used to find the correct correlate, one can hypothesise that emphasising the contrasting words will amplify their prominence in the discourse, facilitating the comprehension of the contrast in the fragmentary answer.

However, it is not only the emphasis that impacts the understanding and acceptability of elliptical sentences but it can be inferred that fragmentary answers are more common in spoken conversations compared to written dialogues due to the disparity between written and spoken language in terms of prioritising complete expressions (cf. Akinnaso, 1982). Given that fragmentary answers are more frequent in spoken language than in written language, one can assume that these structures will be perceived as more acceptable, if they occur in contexts that they are used in more often. Therefore, it is hypothesised that generally, auditory stimuli will receive higher acceptability ratings than written stimuli.

In addition, it can be postulated that not only emphasis and modality have the potential to affect the perception of contrasting fragmented responses, but the semantics of the used words could also play a role in determining how effortlessly the listener or reader establishes a focus-oriented anaphoric connection. That is, the meaning of the words might impact how the fragmentary answers are processed.

Words in language can be categorized into two distinct groups based on their meaning. While lexical words consist of content words such as nouns, verbs, and adjectives, functional words are a closed group, have grammatical meaning, and include, for example, articles, prepositions, and pronouns (cf. Biber et al., 2021). Several studies have shown that lexical words often play a more crucial role in understanding messages. This can be seen in reading as well as in spoken language. In an eye-movement study, Roussel et al. (2018) found that functional words are more likely to be eluded than lexical words and that functional words are fixated less often and with shorter gaze durations than lexical words. In spoken language, stress is usually placed on lexical words (cf. Chomsky & Halle, 1991; Féry & Herbst, 2004).

In the context of this current investigation, participants will be exposed to conversations containing both lexical and functional fragments, with the

aim of examining potential divergences. It is hypothesised that dialogues involving lexical fragments will receive higher acceptability ratings than those incorporating functional fragments.

2.4 Acceptability judgement tasks in linguistics

In this section, the method of acceptability judgement tasks is introduced, encompassing a concise overview of its historical background and previous research. Moreover, the distinction between formal and informal data is explained, along with an elaboration of the commonly employed 7-point scale in these tasks. As a last point, differences between written and auditory stimuli in such judgement tasks are examined, accompanied by the acknowledgement of inherent limitations tied to this methodology.

More than 60 years ago, acceptability judgements were initially suggested as a substitute for assessing grammaticality of syntactic theories. Chomsky (1957) proposed that "[o]ne way to test the adequacy of a grammar proposed for [a language] is to determine whether or not the sequences that it generates are actually grammatical, i.e. acceptable to a native speaker" (p. 13).

Although acceptability judgement tasks (henceforth, AJTs) are employed to examine a language's syntactic structures, the term *grammaticality judgement* is misleading. It assumes that participants possess the ability to consciously access their implicit knowledge about language. Instead, the term *AJT* is used to clarify that insights into the grammaticality of specific syntactic patterns can be gleaned from the acceptability of these structures (cf. Schütze, 2016; Sprouse et al., 2013). In the current study, participants were specifically requested to rate the acceptability of sentences, i.e. how natural they perceive the sentences to be. Therefore, the terms *acceptability judgements* and *naturalness judgements* will be used interchangeably.

Concerning the research involving AJTs, a distinction needs to be drawn between formal and informal approaches. In the following, the term *informal method* refers to AJTs conducted with a low number of participants and are associated with scale biases, judgement errors, etc. In contrast, the term *formal method* pertains to AJTS adhering to experimental standards,

encompassing a common rating scale, a sufficient participant count, and other stipulations (cf. Juzek, 2016).

Employing multi-point scales, exemplified by the implementation of a 7-point Likert scale in this study, offers a versatile means of assessment. While the labelling of the scale can differ across various studies depending on the research question, 1 signifies complete unnaturalness, and conversely, 7 indicates complete naturalness in the present study. As a result, 4 serves to pinpoint structures that neither distinctly lean towards unnaturalness nor naturalness. Thus, 7-point Likert scales allow participants to also identify a balanced midpoint (cf. Sprouse et al., 2013).

Moreover, 7-point Likert scales offer the advantage of enabling statistical analysis of judgements. This includes calculating sample means, standard deviations, and assessing the significance of the effects under investigation (cf. Featherston, 2008). Nonetheless, there are other formal methods such as magnitude estimation, two-alternative forced-choice, and others. For a comprehensive understanding of each method and an in-depth analysis of the appropriate techniques for conducting acceptability judgements, readers are directed to the work of Sprouse et al. (2013).

Although AJTs have traditionally relied on written stimuli, auditory stimuli have been adopted by AJTs as a substitution for written stimuli in recent studies (cf. Jasso, 2022; Kayali, 2023, Liu et al., 2022). This is particularly beneficial when examining structures that are uncommon in written language or necessitate prosodic cues for a comprehensive understanding of the syntactic structure. Therefore, auditory stimuli are deemed suitable in such cases (cf. Sedarous & Namboodiripad, 2020). For an overview of the importance of modality and prosodic emphasis in sentence comprehension, see chapter 2.3.

Juzek (2016) investigated whether the modality of stimuli influences participants' ratings in AJTs. That is, he stated as a null hypothesis that the ratings are the same for both written and auditory stimuli. As an alternative hypothesis, he proposed that constructions that are more common in spoken language receive higher acceptability ratings as auditory stimuli and constructions that are more common in written language receive higher ac-

ceptability ratings as written stimuli. In his experiment, the difference be-tween commonly used written and spoken constructions as written and au-ditory stimuli, respectively, lacked significance. Hence, the null hypothesis could not be rejected. However, the experiment investigated resumptive pronouns. As the present study examines dialogues involving contrastive focus and fragmentary answers, the intonation of those sentences perhaps plays a larger role than for sentences with resumptive pronouns. Therefore, it might be that a significant difference in the modality of stimuli can be found in the present study. For a justification of the present methodology, see chapter 3.1.

As a final aspect, the limitations of AJTs are addressed to provide a comprehensive overview of this methodology. First, AJTs are inherently subjective and hence, leading to a high degree of variations. However, through a sufficient participant count, reliable results are ensured. Second, although a 7-point Likert scale is more fine-grade than binary acceptability judgement, it still may be too general to capture the full range of variation and subtleties in native speakers' intuition. Furthermore, it is worth noting that in some cases, there might be a lack of contextual information, poten-tially hindering the resemblance to natural language use and compromising the reliability of judgements. However, in the current experiment, consider-able efforts were made to select stimuli that minimise the risk of misunder-standings. Additionally, the inclusion of seven stimuli per condition serves to further reduce this risk and enhance the reliability of the results.

3. Data and method

This chapter provides a comprehensive discussion on the study's design along with the process of selecting stimuli, the procedures for recording, the methodologies for data collection, participant recruitment and their attrib-utes, and the techniques employed for data analysis. Furthermore, this chapter delves into the reasoning, methodologies, and factors considered for each of these components.

3.1 Study design

The experiment was conducted using a 2 (modality: written or auditory) x 2 (emphasis: with or without emphasis) x 2 (fragment-type: functional or lexical word) study design. Hence, eight conditions were tested by using three binary factors. It is important to note that the research was conducted as an online study, and the participants were sourced from the platform *Prolific*. For an in-depth explanation of the participant recruitment, see chapter 3.5.

A between-subject design was employed to examine the effects of modality, while a within-subject design was utilised to investigate the influence of emphasis and fragment-type. That is, participants were randomly divided into groups receiving either written or auditory stimuli exclusively. However, each participant rated stimuli featuring different emphasis levels and types of fragments. The chosen study design aims to mitigate participant perplexity or scepticism arising from varying modalities and to ensure that any observed differences in results for stimuli with different emphasis and fragment types are attributable to their influencing factors rather than participant variability. This design selection safeguards against potential confounding factors, reducing individual differences and increasing the sensitivity to detect effects.

The study design is based on two short pilot studies, which were completed by a total of 18 participants. The aim of the pilot studies was to ascertain the clarity of instructions presented in the introductory pages of the experiment and to confirm the methods of the study itself, including aspects such as the audio files and the randomised grouping process.

In the study, participants were asked to rate dialogues regarding their naturalness. The AJT was an ordinal response task on a 7-point Likert scale. The 7-point Likert scale used in the present study asked participants to rate each dialogue presented in the experiment on a scale from 1 to 7, representing varying degrees of acceptability. That is, the scale ranged from 1, representing fully unnatural structures to 7, indicating full naturalness. Prior, fully natural, fully unnatural, and neither natural nor unnatural examples were given in the introductory part to the study that are explained in more detail in the following section.

After a welcoming page, participants were presented with three dialogues (13-15) that had a similar structure to the critical and filler items of the experiment and varying acceptability ratings. The dialogues did not include the variables that were investigated in the experiment. However, through the introduction of similar dialogues, participants became familiar with the rating scale. Moreover, it was ensured that all participants understood that the study aims to determine what sentences would be acceptable in daily speech contrary to written language (cf. Sedarous & Namboodiripad, 2020). Note that the glossing and translation is given in (13-15), while participants were only presented with the German sentences.

(13) A: Was mag Peter?
what likes Peter
'What does Peter like?'
B: Peter   mag Ingwer.
Peter   likes ginger
'Peter likes ginger.'
(adapted from Sedarous & Namboodiripad, 2020, p.7)

(14) A: Was hat Peter   gestern   gemacht?
what AUX Peter   yesterday did
'What did Peter do yesterday?'
B: *Vater   Fußball   gestern.
father   football   yesterday
'father football yesterday.'

(ibid.)

(15) A: Hat Peter   inzwischen   aufgegeben?
AUX Peter   by.now   gave.up
'Has Peter given up by now?'
B: ?Nein,   das Handtuch,   das würde   er
no   the towel   that AUX   he
bestimmt nie   werfen!
certainly never   throw
'No, the towel, he would certainly never throw that in!'
(adapted from Wierzba et al., 2023, p. 16)

Participants were instructed that their acceptability ratings should be based on only speaker B's response to speaker A's utterance. While (13) is described to the participants as fully natural, (14) is identified as fully unnatural and (15) is used as an example of an utterance that is neither natural nor unnatural.

After the introductory pages, participants were randomly assigned to either only written or auditory stimuli. Both conditions included seven items of each variable, i.e. with and without emphasis of contrasting words as well as functional and lexical fragments. In total, 56 critical items, including both written and auditory items, were used for the study. However, since each participant was assigned to either written or auditory stimuli, each participants encountered 28 critical items and 28 filler items. The critical, written items were equally distributed across the four conditions, i.e. with and without orthographic marking as well as lexical or functional fragment types. In the written condition, participants were informed that capitalisation within the dialogues serves to indicate which words are emphasized by the speaker. Similarly, the critical auditory items were equally distributed across the four conditions, i.e. with and without prosodic marking on the contrasting words as well as lexical or functional fragment types. Therefore, each condition is exemplified by seven items in each run of the experiment. This balanced design allows for a systematic examination of the effects of emphasis and fragment types on the experimental variables.

The study was designed in a way that prevented participants from revisiting previous slides and changing their ratings or skipping dialogues without providing their acceptability ratings. Additionally, participants in the study involving auditory stimuli were required to listen to the entire audio files of the dialogues before they could give their ratings.

At the end of the questionnaire, participants were asked to voluntarily indicate their age, gender, level of education, and native language or native languages. Moreover, participants could leave additional comments. Completing the entire questionnaire took the participants approximately 10 minutes.

This methodological choice was motivated by several factors. First, the 7-point Likert scale offers an appropriate range of response options, allowing participants to express nuanced judgements effectively. It includes a balanced midpoint that signifies structures perceived as neither natural nor unnatural, as discussed in chapter on 2.4. Second, the inclusion of auditory stimuli enables the capture of the full range of linguistic cues present in natural speech of spoken languages. Dialogues, particularly those involving

contrastive focus and fragmentary answers, often rely on prosodic features, as discussed in more detail in 2.5. By introducing participants to auditory stimuli, a more ecologically valid depiction of these linguistic cues is offered in contrast to relying solely on written stimuli. Third, including written stimuli alongside auditory ones enables the exploration of potential differences or convergences in acceptability judgements between the auditory and written presentations, shedding light on the role of modality in the perception of naturalness, specifically for fragments and contrastive focus.

3.2 Selection of critical and filler items

This section gives an overview of the selected critical and filler items for the study and the rationale behind their choice. The list of written critical and filler items can be found in the appendix, while their verbal equivalents can be found on this page: https://shorturl.at/anqsD.

An exemplary overview of how the written and auditory stimuli vary regarding emphasis and fragment type is shown in (1-4), repeated here as (16-19).

(16)   A: Peter hat AB 18 Uhr im Kino gearbeitet.
              'Peter worked at the cinema FROM 6pm.'
        B: Nein, BIS 18 Uhr.
              'No, UNTIL 6pm.'
                                                                    (own example)

(17)   A: Peter hat ab 18 Uhr im Kino gearbeitet.
              'Peter worked at the cinema from 6pm.'
        B: Nein, bis 18 Uhr.
              'No, until 6pm.'
                                                                    (own example)

(18)   A: Peter hat dem POLIZISTEN seinen Ausweis gezeigt.
              'Peter showed his identity card to the POLICE OFFICER.'
        B: Nein, dem TÜRSTEHER.
              'No, the BOUNCER.'
                                                                    (own example)

(19)   A: Peter hat dem Polizisten seinen Ausweis gezeigt.
              'Peter showed his identity card to the police officer.'
        B: Nein, dem Türsteher.
              'No, the bouncer.'
                                                                    (own example)

The stimuli in (16) and (17) include functional fragments, namely prepositions, while the stimuli in (18) and (19) incorporate lexical fragments, namely nouns denoting human referents. Moreover, the stimuli in (16) and (18) emphasise the contrasting words and their verbal equivalents bear pitch accent on said words. In contrast, the stimuli in (17) and (19) do not incorporate any orthographic marking and their verbal equivalents display natural intonation of the sentence.

Several steps were taken to minimise the influence of extraneous factors. First, the contrasting words in the stimuli with functional fragment type incorporated the prepositions *bis* 'until' and *ab* 'from', *mit* 'with' and *ohne* 'without', as well as *nach* 'after' and *vor* 'before', as these demonstrate opposite meanings.

Second, for the lexical fragments and their correlates, the contrasting nouns all denoted human referents. Moreover, only masculine nouns marked overtly with dative case were chosen to stand in contrastive focus to ensure that the reader or hearer can unambiguously identify the correlate of the fragment, as explained in chapter 2.3.

Third, critical items were adjusted to be in past tense to ensure that the word in contrastive focus is not in final position. Moreover, to allow for unbiased research on contrastive focus, only the indirect object that precedes the direct object, was chosen for the contrastive focus in the allow for unbiased research on contrastive focus in the condition with lexical fragments, as outlined in chapter 2.3. Similarly, in the condition with functional fragments, the prepositional phrase precedes the object phrase to ensure a consistent pattern.

Next, the sentences were created in such a way that stimuli with lexical fragments include ditransitive verbs, while stimuli with functional fragments, which must include a preposition phrase based on the study design, only include transitive verbs, as can be seen in the example (18-21). Therefore, a comparatively equal length of all stimuli is guaranteed.

Last, the contrasting words are either orthographically or prosodically marked in the condition with emphasis on the one hand. In the condition without emphasis, on the other hand, the stimuli either do not contain any orthographic marking or the nuclear accent is not on the contrasting word

but on the default position (cf. Féry, 2011). The orthographic marking involved writing the respective words in uppercase letters. For a detailed account and a graphical representation of the pitch contour for stimuli with prosodic marking and those with default intonation, see chapter 3.3.

A total of 56 critical items were selected for the study. However, due to the study's design, each participant only encountered 28 critical items that were either only written or only auditory stimuli. In addition, 56 filler items were included in the study, of which 28 were written and 28 were auditory items.

Filler items involved dialogues that incorporated either non-fragmental contrast such as in (20) or dialogues without any contrast such as (21)[1]. The order of critical and filler items was randomly arranged, with each item being presented on its own individual page.

(20)   A: Peter   hat   die   SÜDDEUTSCHE   gelesen.
               Peter   AUX   the   Süddeutsche   read
               'Peter read the Süddeutsche.'
         B: Nein,   er   hat   die   FAZ gelesen.
               no   he   AUX the   FAZ read
               'No, he read the FAZ.'

(21)   A: Peter   hat   in   der   Mensa   zu   Mittag
               Peter   AUX   in   the   canteen   for   lunch
               gegessen.
               ate
               'Peter had lunch in the canteen.'
         B: Ja,   zusammen   mit   Freunden.
               yes   together   with   friends
               'Yes, together with friends.'

The acceptability of the fillers varied. The fillers in (20) and (21) above represent structures associated with full naturalness. The following fillers in (22) and (23) signify complete unnaturalness.

(22)   A: Peter   hat   mit   Freunden UNO   gespielt.
               Peter   AUX   with   friends   UNO   played
               'Peter played UNO with friends.'
         B: *Nein,   beim   Stammtisch   die   Freunde

---

no   at.the   regulars' table the   friends
haben   mit Vorliebe   SKAT   gespielt.
AUX   with preference   Skat   played
'No, at the regular's table the friends played skat with preference.'

(23)   A: Peter   hat   seinem   Sohn   ein   Geschenk
              Peter   AUX his   son   a   gift
              gemacht.
              made
              'Peter gave a gift to his son.'

         B: *Ja, ein Fahrrad in   die   Schule   zum Fahren.
              yes, a   bike   to   the   school   for riding
              'Yes, a bike to the school for riding.'

To ensure consistency, multiple steps were taken. First, all stimuli, i.e. critical and filler items, were adjusted to be in past tense and start with *Peter*. Second, approximately half of the filler items incorporated orthographic or prosodic marking on the contrasting words, while the remaining half lacked such marking, mirroring the divergence seen in the critical items. Next, out of the 56 filler items, 10 items represented full acceptability, 12 items indicated some acceptability, 12 items denoted neutrality in terms of acceptability, 12 items implied partial unacceptability, and 10 items signified full unacceptability. This approach ensured an equivalent count of natural and unnatural supplementary items, and the spectrum of acceptability reflects that assumed for the critical items.

3.3 Recording of the stimuli

This section is concerned with the process of recording stimuli, discussing the methodologies and considerations involved in capturing high-quality audio or visual materials for the present research study.

Stimuli were recorded in the open-source toolkit *Praat* in a soundproof room, using a Blue Snowball ICE microphone and saved to be in WAV-format. Silences before and after the sentences were cut out of the sound files. As each stimulus represents a dialogue, the two parts had to be recorded individually.

Each part of every stimulus was recorded at least three times, of which the one with the highest clarity, intelligibility, and adherence was chosen for the experiment. The first part was recorded by the voice actor Roman Pertl

(henceforth, speaker A). The second part was recorded by the author (henceforth, speaker B). Both speakers are native German speakers, were familiar with the sentence prior to the recording, and had the opportunity to re-record any sentence indefinite times.

In accordance with the guidance provided by Sederous and Namboodiripad (2020), sentences were organised by condition, recorded multiple times, and deliberately excluded any instances of inhalation or exhalation sounds. The pitch contour of the stimuli (20) and (21) in chapter 3.2 are illustrated in Figure 1 and 2.
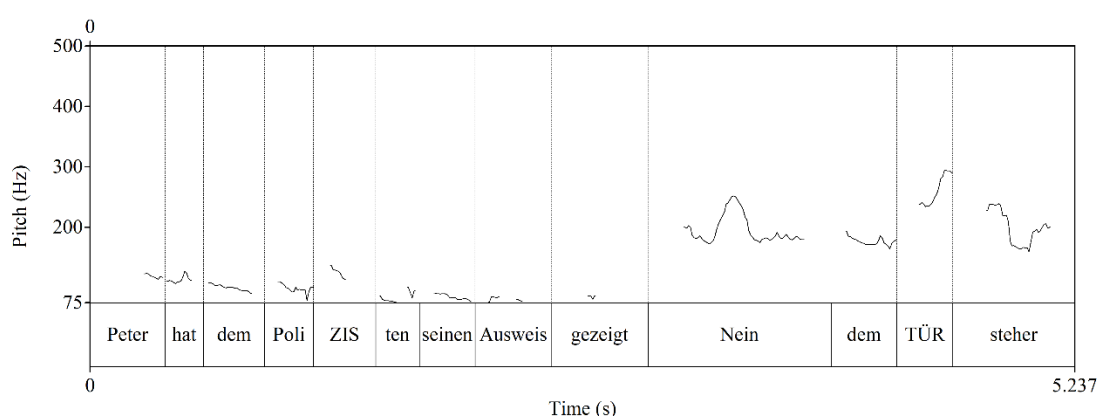
Figure 1: Pitch contour of stimulus with emphasis

Figure 1 shows the recorded intonation of the stimulus (20), that includes the emphasis of the contrasting words *Polizisten* 'police officer' and

Türsteher 'bouncer'. Both words are marked with L+H* accent. The intonational contour of the preceding sentence in Figure 1 stands in stark contrast with its equivalent in Figure 2.
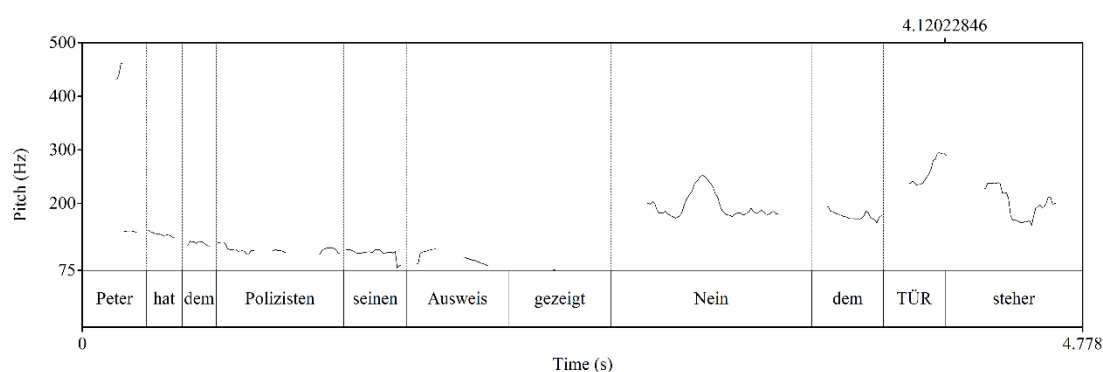

Figure 2: Pitch contour of stimuli without emphasis

Speaker A's sentence in Figure 2 shows consistent, natural intonation. That is, the preverbal position, i.e. on the word *Ausweis* 'identity card', represents the default sentence accent (cf. Féry, 2011). As becomes apparent in the pitch contours, the stimuli differ in whether the word *Polizisten* 'police officer' is emphasised or not. While *Polizisten* in the stimulus displayed in Figure 1 received L+H* accent, it is de-accented in the stimulus in Figure 2. Note that the apparent distinction in pitch between the speakers is attributable to the gender contrast (cf. Simpson 2009), with the first speaker being male and the second speaker being female.

Speaker B's recording, that places L+H* accent on the contrasting word *Türsteher* 'bouncer', was used for both conditions. That is, when combining the parts of speaker A and speaker B, the same recording of speaker B's answer was used for the stimuli in the conditions with and without emphasis in the preceding sentence to ensure consistency and minimise confounding factors. Given that the contents of the sentences as well as the fragmentary answers were identical, the stimuli depicted in Figure 1 and Figure 2 solely vary in terms of the emphasis placed on *Polizisten* 'police officer'. This deliberate difference serves to eliminate alternative explanations for the observed outcomes, strengthening the validity of the results.

After the recording, the audio files of each speaker were concatenated in Praat. Therefore, each stimulus is composed of two merged audio files,

seamlessly transitioning from the first to the second audio file without any audible disruptions or breaks.

Next, the audio files were controlled for loudness in Praat using a plugin (https://www.praatvocaltoolkit.com/normalize.html). The raw recordings as well as the combined, neutralised recordings can be found on this page: https://shorturl.at/fiDLT.

3.4 Data collection

In this chapter, the data collection process employed in the research study, which involved gathering acceptability judgements from participants, is discussed. Moreover, the methods employed to obtain these judgements, encompassing the experimental design and data collection procedures, are explained.

Acceptability judgements were crowdsourced from *Prolific* (cf. Prolific Academic, 2019). Prolific serves as an online platform employed to recruit participants for research studies. It ensures a high level of transparency for both researchers and participants, features a user-friendly interface, and enables the prescreening of potential participants (cf. Palan & Schitter, 2018).

In the experiment, stimuli were presented either as text or audio file. In the auditory scenario, participants were able to provide their ratings only after they had finished listening to the audio file. In both conditions, participants were required to assign a rating to the current dialogue before proceeding to the subsequent one, and in the presentation occurred one dialogue at a time. Participants were not permitted to revisit prior dialogues to modify their ratings. There were no time constraints in place. That is, participants were allowed to take as much time as necessary for reading the dialogues and potentially re-listening to the audio files.

When presented with the dialogues, participants were asked the following question in the experiment. Note that translation is given in (24), while participants were only presented with the German question.

(24)   Wie natürlich wirkt die Aussage der Sprecherin B auf Sie?
        'How natural does speaker B's response sound?'
                            (adapted from Featherston, 2008, p. 6)

The question of naturalness in (24) was taken from other studies incorporating AJTs (cf. Featherston, 2008). However, since all participants were native German speakers, the question was translated to German. Moreover, the question was adapted in such a way that it includes *auf Sie* 'to you', asking for participants' individual unconscious knowledge of the German language. Therefore, it is ensured that participants do not rate the naturalness of the presented dialogues based on prescriptive grammar rules but their own intuition as a native speaker.

Due to the setup of the study, it was ensured that participants were randomly assigned to different conditions and although participants were grouped into either the written or auditory condition, they were unaware that they remained uninformed about the existence of alternate conditions involving different media, thereby effectively minimising the potential for biases.

3.5 Participant recruitment and characteristics

The data for this study was obtained from *Prolific*, with a minimum approval rate requirement of 90% to ensure reliable work. Additionally, the study was carefully setup to exclusively include participants who self-identified as native German speakers, as the present study focuses solely on the German language.

The study was completed by a total of 100 participants, comprising 69 males, 29 females, and 2 individuals who identified as diverse. The participants' ages varied between 19 and 73 years, with a mean age of 35.53 years. Among them, 9 participants did not hold a high school diploma, 29 participants completed high school as their highest level of education, 27 participants attained a bachelor's degree, and 35 participants had a higher degree beyond the bachelor's level.

Before the commencement of the study, participants were acquainted with the voluntary nature of their involvement and the exclusive utilisation of their data for scientific research purposes. Stringent measures were taken to uphold the privacy and confidentiality of participants' data. Owing to the study's design, the socio-demographic details provided could not be linked

to individual identities. Every participant was allowed to participate only once and was paid for their participation.

4. Results and analysis

The following sections provides an overview of the study's findings. First, the raw data from the questionnaire, comprising participants' responses, is presented. Second, the methods used for data analysis are explained in detail. Next, the results of the data analysis are shown, followed by the addressing of the hypotheses. This comprehensive approach will help clarify the study's outcomes and shed light on the obtained results.

4.1 Descriptive statistics

The study aimed to investigate to what extend the emphasis, modality, and fragment type in dialogues involving contrastive focus and fragmentary answers affect how native German speakers perceive the naturalness of such contrastive fragmentary answers. This section presents participants' responses from the questionnaire, providing an unprocessed view of the collected data.

Out of the 100 participants who completed the study, 57 participants were presented with auditory stimuli, while 43 participants were faced with written stimuli. The somewhat unequal distribution between these two conditions is a result of the participants being randomly assigned. Due to the setup of the study, all participants encountered stimuli with varying levels of emphasis and varying fragment types. Except for removing filler items, no data points were eliminated from the dataset. The analysis incorporated the complete dataset obtained from the experiment's critical items.

First, consider Figure 3, which shows participants' responses to all critical items. The graph displays participants' acceptability ratings based on the varying conditions of the three investigated factors. The ratings to stimuli with emphasis are illustrated in the graphs on the top, while the graphs on the bottom present participants' ratings to stimuli without emphasis, encompassing both conditions for the investigated factor emphasis. Similarly, ratings to auditory stimuli are illustrated on the left, while the rating for written

stimuli is shown on the right. Moreover, functional fragments are shown in red, whereas lexical fragments are coloured in blue.
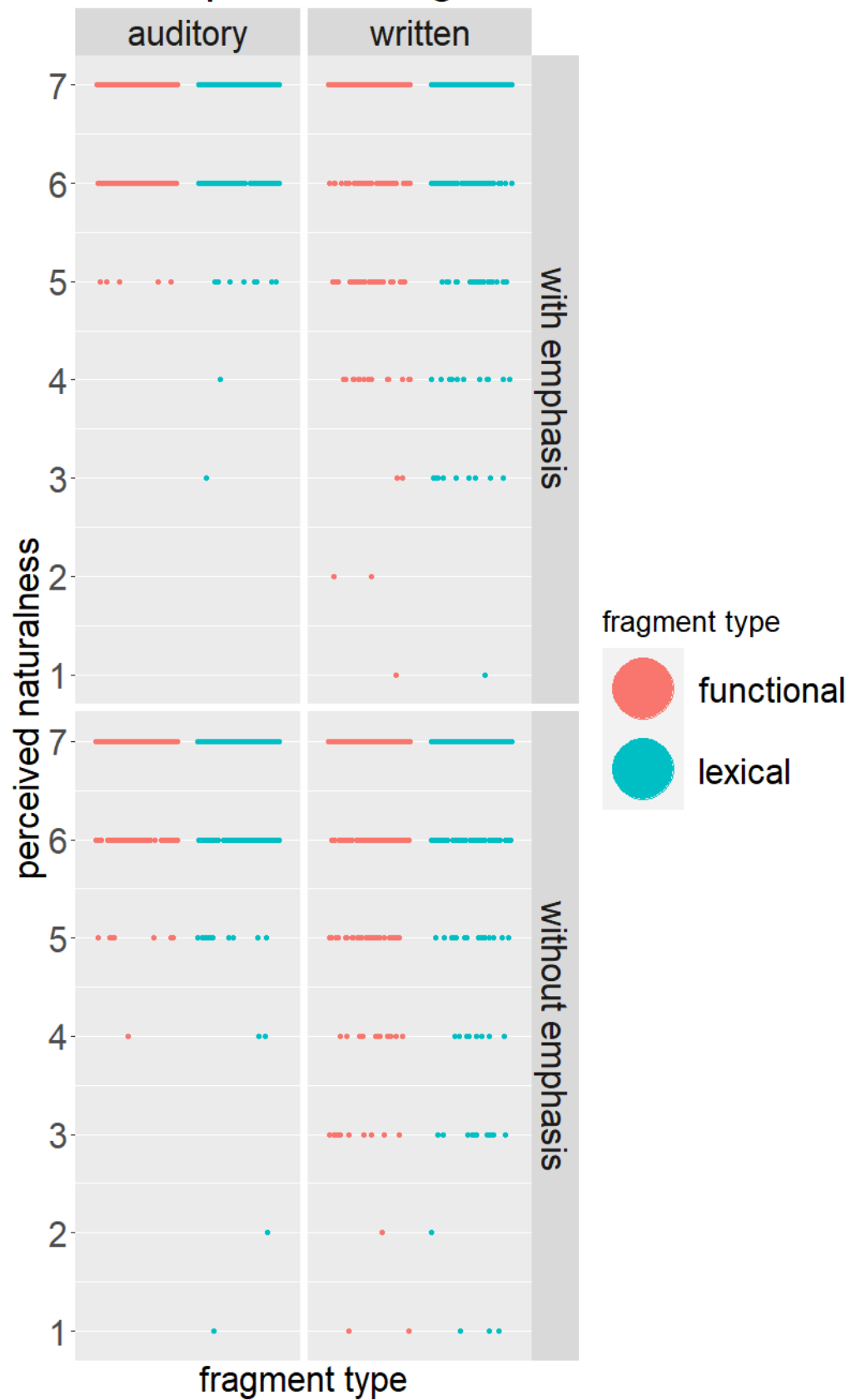
Participants' ratings of all critical items

Figure 3: Scatter plot of participants' ratings of all critical items

As evident in Figure 3, a substantial number of responses gravitated towards the rating of 7 (fully natural). Figure 3 serves as a preliminary overview. In the remaining of this chapter, the analysis will delve deeper, thoroughly examining the means and variances associated with each factor.

Now, consider Figure 4, displaying participants' ratings for each investigated factor in mosaic plots.

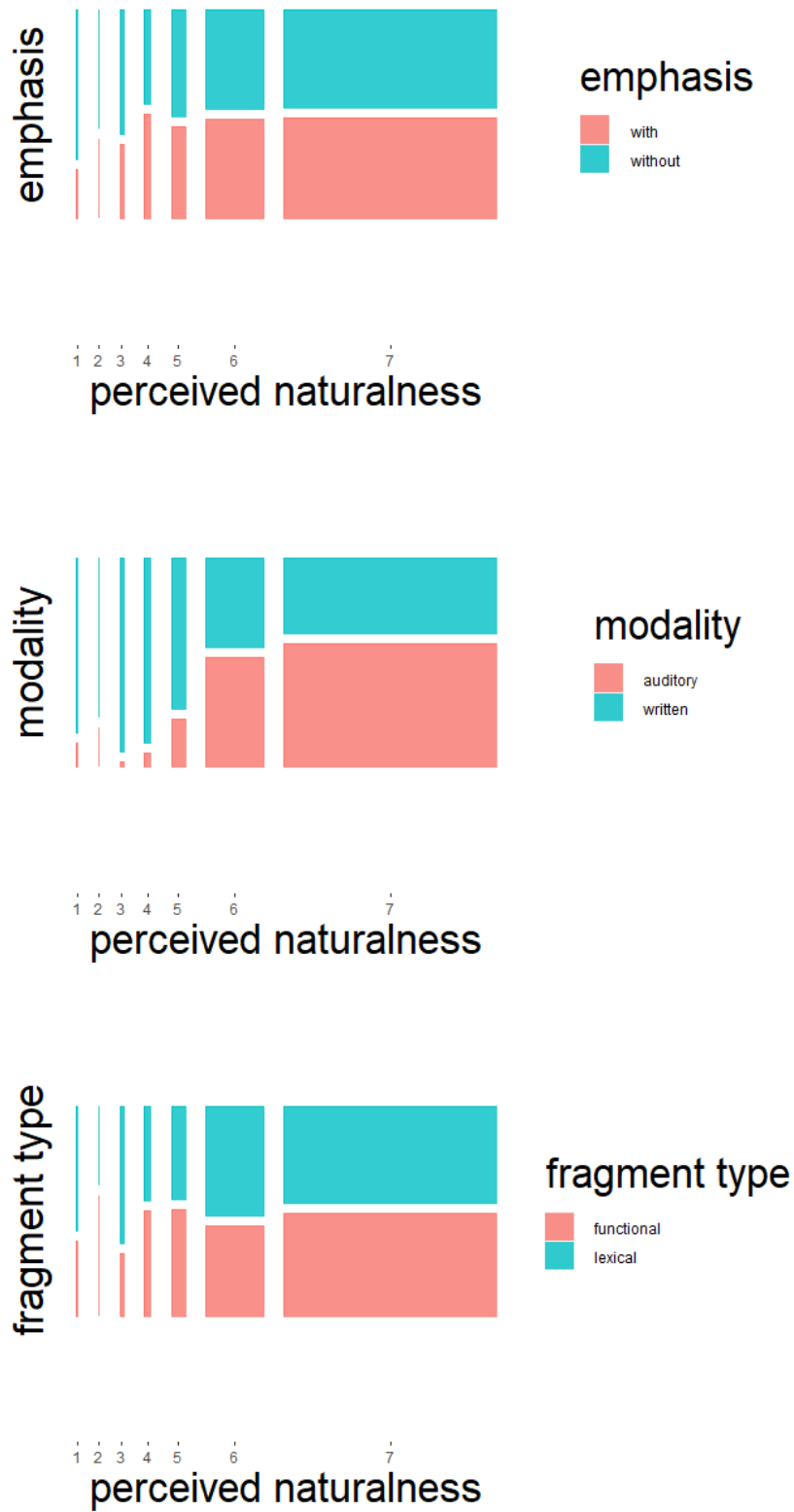# Participants' ratings shown for each factor

Figure 4: Mosaic plots of participants' ratings of each factor

The trend which has already been visible in Figure 3 becomes more apparent in Figure 4. Specifically, most of the fragmentary responses received a rating of 7 (fully natural) on the 7-point Likert scale, independent of the experimental condition.

On the one hand, emphasis and fragment type exhibit remarkable similarity in the higher ratings. Nevertheless, the mosaic plot reveals marginal disparities in the lower ratings. On the other hand, when considering modality, a noticeable differentiation is already apparent between ratings 6 and 7, with this contrast being more pronounced in the lower ratings. In this context, the examined variables appear to exert a more profound influence on participants' ratings. This divergence underscores the influence of the investigated factors on participants' judgements of acceptability, highlighting variations in how distinct factors shape their perceptions.

The first hypothesis examines whether the inclusion of emphasis on contrasting words impacts the perceived naturalness of contrastive, fragmentary answers. Therefore, consider Figure 5, which illustrates participants' ratings of stimuli with and without emphasis.
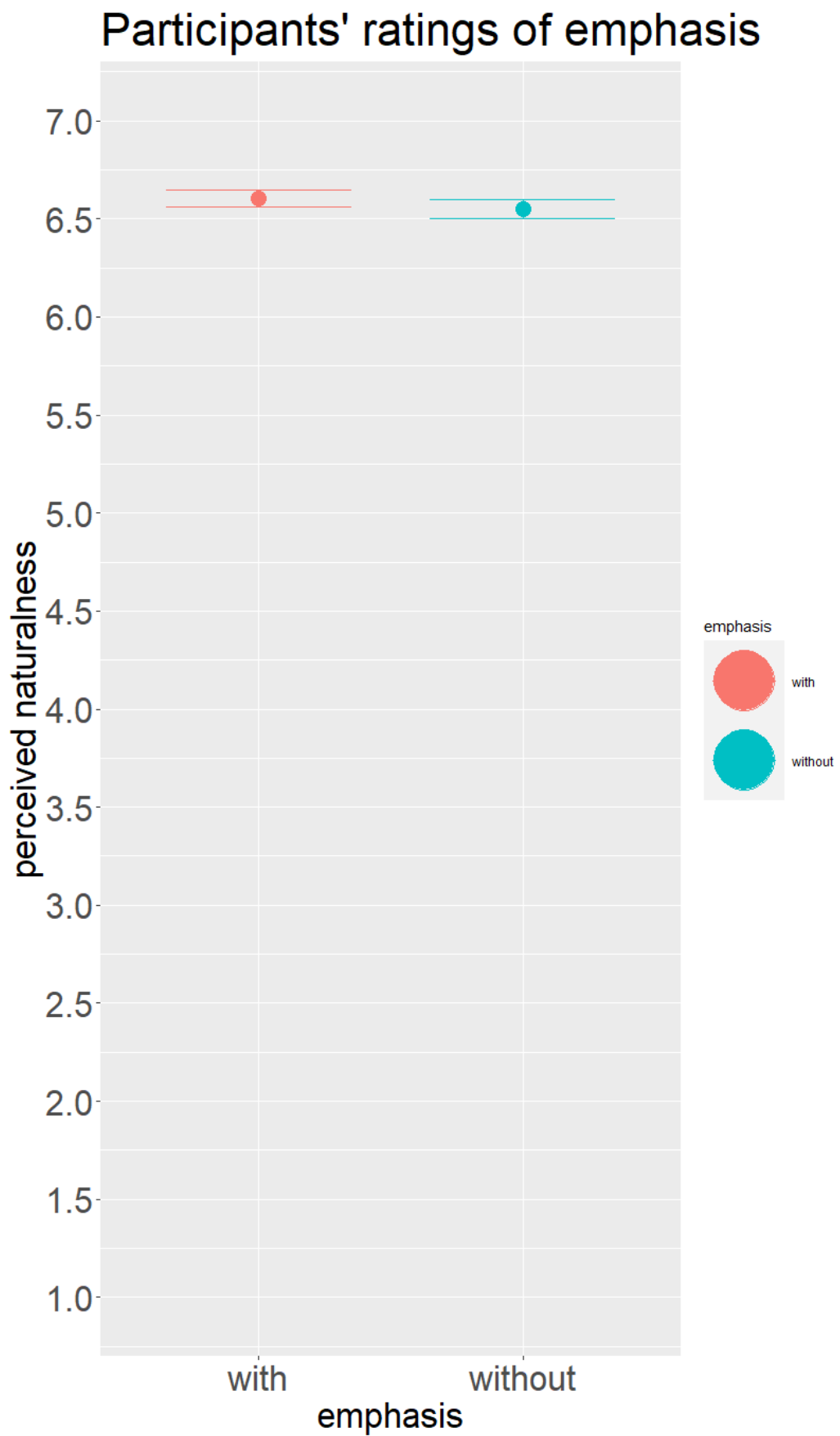
Figure 5: Comparison of participants' ratings of stimuli with and without emphasis

As depicted in Figure 5, a subtle distinction emerges in the ratings given to fragmentary responses. In fact, participants assigned a rating of m = 6.60 with a standard deviation of sd = 0.80 on the 7-point Likert scale to fragmentary answers with emphasis, whereas fragmentary answers lacking emphasis received a slightly lower rating of m = 6.55 with a standard deviation of sd = 0.91.

Next, consider Figure 6, which shows the mean ratings of auditive and written stimuli. Participants' ratings of auditory stimuli are illustrated in red, whereas those of written stimuli are coloured in blue. According to the second hypothesis, auditory stimuli would receive higher acceptability ratings than their written counterparts.
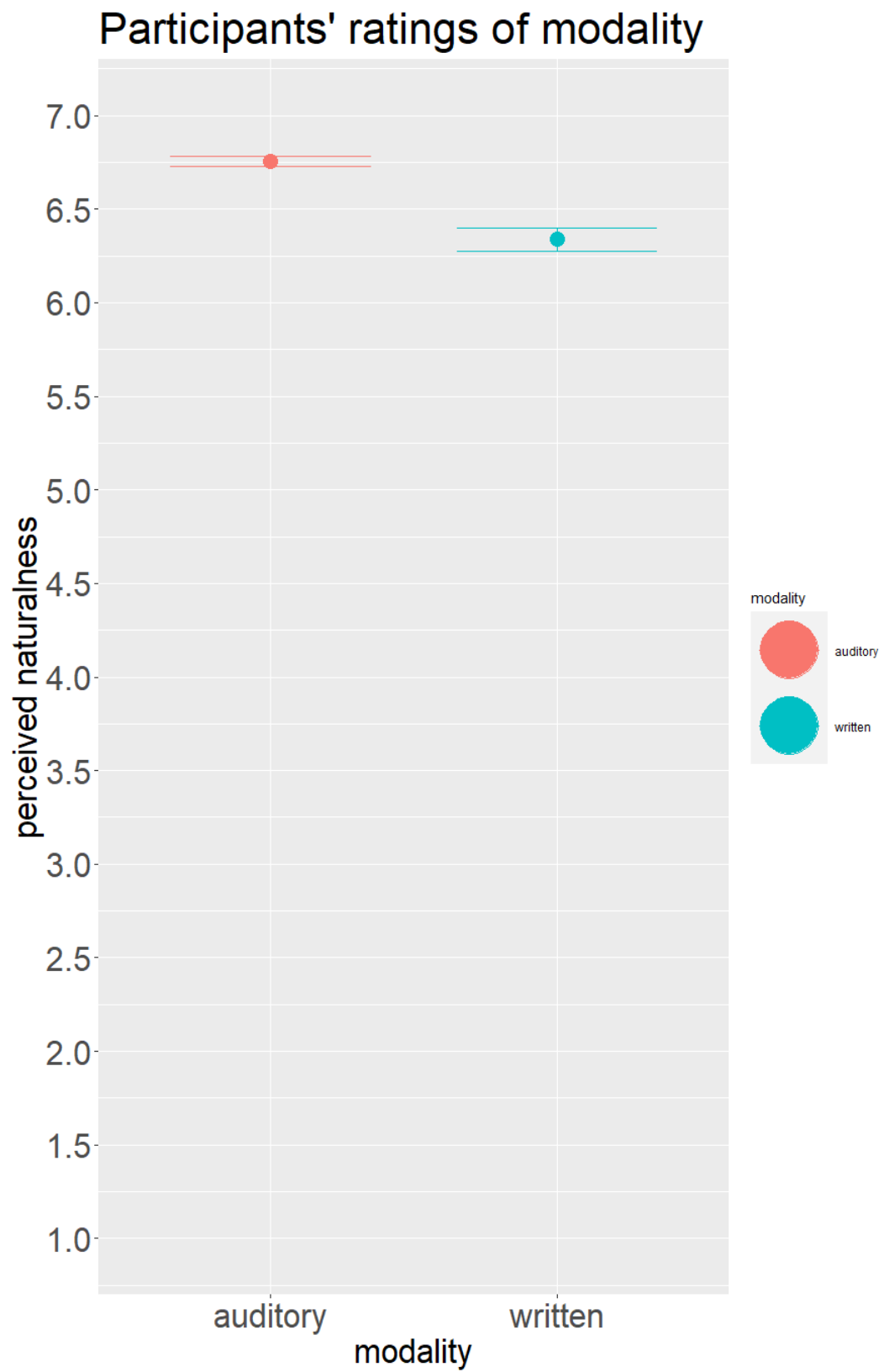
Figure 6: Comparison of participants' ratings of auditory and written stimuli

As depicted in Figure 6, a notable contrast can be observed between the ratings in the auditory condition as opposed to those in the written condition. That is, fragmentary answers in the auditory condition received a rating of $m = 6.76$ with a standard deviation of $sd = 0.53$ on the 7-point Likert scale, whereas fragmentary answers in the written condition received a slightly lower rating of $m = 6.34$ with a standard deviation of $sd = 1.11$.

Last, the third hypothesis explores whether fragment type influences the perception of naturalness in fragmentary answers. As illustrated in Figure 7, there is a subtle contrast in acceptability ratings between functional and lexical fragments. In fact, participants assigned a rating of $m = 6.61$ with a standard deviation of $sd = 0.82$ on the 7-point Likert scale to answers incorporating functional fragments, whereas those with lexical fragments received a slightly lower rating of $m = 6.55$ with a standard deviation of $sd = 0.90$.
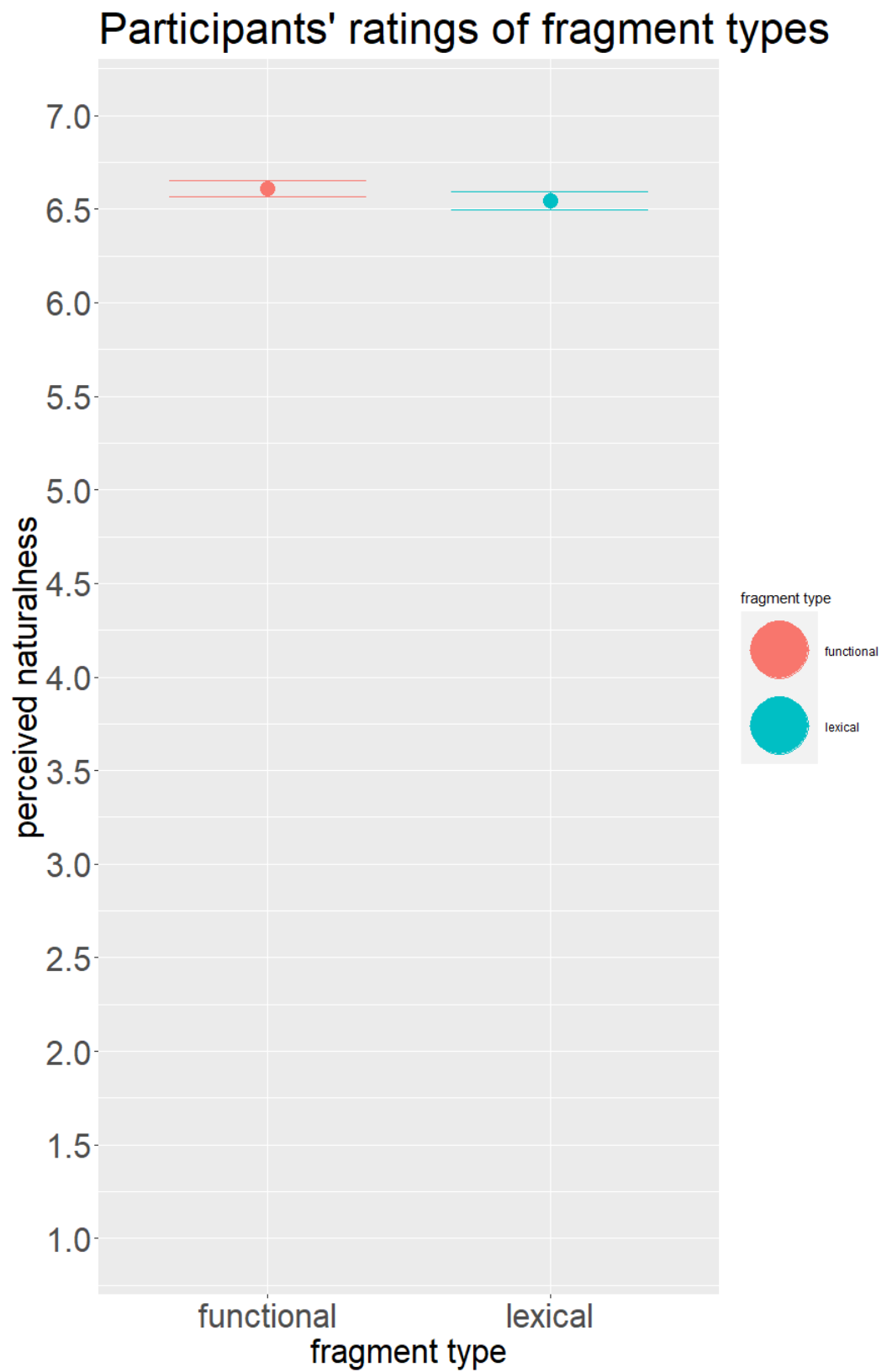
Figure 7: Comparison of participants' ratings of stimuli with functional and lexical fragments

4.2 Methods of analysis

This section provides an in-depth explanation of the methods chosen for data analysis, along with the rationale behind their selection.

Although the descriptive statistics presented in Chapter 4.1 provided a general overview, a more detailed presentation of the data and a statistical analysis is required to answer the study's research questions. Therefore, participants' Likert scale responses were z-scored and analysed using Cumulative Link Mixed Models (henceforth, CLMM), using R version 4.2.1 (cf. R Development Core Team, 2015).

In order to facilitate meaningful comparisons and analyses, the 7-point Likert scale ratings provided by participants were standardised through a z-scoring procedure. Z-scoring involves subtracting the mean rating across all participants from each individual rating and then dividing by the standard deviation of the ratings. Z-scoring enhances the interpretability and comparability of the ratings across different factors and conditions, enabling a more comprehensive exploration of the underlying patterns and effects (cf. Curtis et al., 2016).

The z-scored 7-point Likert scale ratings were analysed using CLMM (cf. Christensen, 2011). CLMM was chosen as the statistical approach due to its suitability for the nature of the data. The CLMM methodology effectively accommodates ordinal responses, making it a robust choice for analysing Likert scale ratings (cf. Taylor et al. 2022). By accounting for the ordinal structure of the data, CLMM captures the inherent order and spacing between the response categories, providing a more accurate representation of participants' perceptions. CLMM models take into consideration both fixed and random effects, allowing to examine the impact of various predictor variables on the odds of participants choosing higher or lower response categories on the Likert scale. This approach is particularly advantageous when investigating factors that may influence participants' perceived naturalness in different conditions or contexts. The use of CLMM acknowledges the inherent correlations within the Likert scale ratings and provides a comprehensive understanding of the underlying relationships between the investigated factors and participants' responses (ibid.). Overall, the application of CLMM aligns with the nature of the data and research objectives,

offering a robust and tailored framework for exploring the effects of different factors on participants' ratings in the study.

To guarantee a robust data analysis, it was assessed whether adding predictors to the model significantly improves the fit compared to a simpler CLMM model with only random intercepts for each participant. That is, the simpler model, so-called null model, was fit to the data and the Akaike Information Criterion (AIC) of this model was computed. AIC is a measure of the model's quality of fit while penalising for the number of parameters. Lower AIC values suggest better model fit. AIC takes into account both how well the model fits the data and the number of parameters in the model. The goal is to find a model that adequately explains the data while avoiding overfitting, which occurs when a model is too complex and captures noise in the data rather than genuine patterns (cf. Cavanaugh & Neath 2018). Furthermore, to ensure a reliable data analysis, the common significance threshold of $\alpha < 0.05$ was adopted (cf. Hedderich & Sachs, 2016).

Last, an analysis of variance (ANOVA) between the null model and the CLMM model used in the data analysis above was conducted to test whether adding predictors to the model significantly improves the fit. ANOVA is a statistical technique used for hypothesis testing. It assesses whether variations in the data can be attributed to differences between groups or if they are simply due to random chance. ANOVA provides an F-statistic and a p-value, with a low p-value indicating that at least one group differs significantly from the others, making ANOVA a valuable tool for comparing multiple groups in research and experimentation (cf. Ståhle & Wold, 1989).

4.3 Results and hypotheses examination

In the following, the data undergoes analysis to address each of the three hypotheses that were the focus of the present study. The code for the analysis can be found on this page: https://shorturl.at/rwV57.

The first hypothesis examines whether the inclusion of emphasis on contrasting words impacts the perceived naturalness of contrastive, fragmentary answers. As already mentioned in chapter 4.1, fragmentary answers with emphasis received a higher rating than those lacking emphasis.

This difference aligns with the predictions of the first hypothesis, suggesting that acceptabili ratings are higher for stimuli with emphasis compared to those without. Employing a CLMM, which treats potential variations among different participants and items as random effects (cf. Baayen et al., 2008), the observed difference of emphasis is statistically significant with a p-value of p = 0.0268, adhering to the predetermined α < 0.05.

Furthermore, the coefficient *without emphasis* in the CLMM model is estimated to be β_1 = -0.2519. This coefficient represents the effect of the absence of emphasis on the perceived naturalness of fragmentary answers. Specifically, a negative coefficient such as β_1 = -0.2519 suggests that when emphasis is not present, the odds of perceiving a fragmentary answer as more natural are reduced compared to stimuli with emphasis. In other words, native speakers tend to find stimuli with emphasis more natural. Therefore, the present data provides evidence in favour of the first hypothesis, indicating that stimuli with emphasis are perceived as more natural by native speakers.

In the present experiment, contrasting words were emphasised either by orthographic marking in the written condition or by prosodic marking in the auditory condition. The present study does not only investigate the influence of emphasis on the perceived naturalness of fragmentary answers, but also delves into how the modality of presentation could influence how native speakers evaluate such fragmentary answers in terms of naturalness.

Hence, the second hypothesis analyses whether modality has an impact on how natural fragmentary answers are perceived. As predicted by the second hypothesis, acceptability ratings are higher for auditory stimuli compared to their written counterparts. In the CLMM, the observed difference of modality holds statistical significance with a p-value of p = 0.0158, adhering to the predetermined α < 0.05.

Furthermore, the coefficient *written* in the CLMM model is estimated to be β_1 = -1.9905. This coefficient represents the effect of written stimuli on the perceived naturalness of fragmentary answers. That is, the odds of perceiving a written contrastive fragment as natural are reduced compared to auditory stimuli. Therefore, the present data provides evidence to reject the

null hypothesis and accept the second hypothesis that modality influences the perceived naturalness of fragmentary answers.

Regarding the final hypothesis, it was assumed that fragmentary answers incorporating lexical words would receive higher acceptability ratings than those including functional words. However, surprisingly, the observed difference in acceptability ratings does not coincide with the predictions of the third hypothesis. Instead, as shown in the present data, the ratings exhibit an inverse trend. That is, dialogues incorporating functional fragments are perceived as more natural than those containing lexical fragments.

In the CLMM, the observed contrast in fragment type is statistically significant with a p-value of $p < 0.01$, adhering to the predetermined $\alpha < 0.05$. The coefficient *lexical fragments* in the CLMM model is estimated to be $\beta_1 = -0.4486$. This coefficient represents the effect of stimuli with lexical fragments on the perceived naturalness of fragmentary answers. That is, the odds of perceiving a written contrastive fragment that incorporates lexical fragments as natural are reduced compared to those including functional fragments. While the present data does not provide evidence in favour of the third hypothesis, it indicates that the difference in fragment type is statistically significant.

The AIC value of the model including emphasis, modality, and fragment type as predictors as well as random effects is 2836.293. On the contrary, the AIC value of the null model which includes only the random effect for submission ID and no predictors is 2864.649. A lower AIC value suggests that the model incorporating emphasis, modality, and fragment type as predictors is better suited to describe the given data compared to the null model. The model with all predictors has found a better trade-off between model complexity and fit to the data. The higher AIC value of the null model indicates that this model is less suited to explain the data compared to the model with all predictors. The null model is less capable of explaining the variation in the data since it lacks predictors.

The difference between the AIC values of the two models is -28. As a lower AIC value indicates a better model fit, this negative difference indicates that the model with all predictors provides a significantly better fit to the data compared to the null model. However, it might be that there is an

effect that does not appear in the model or that predictors are rather weakly significant.

The result of the ANOVA analysis indicates that using emphasis, modality, and fragment type as predictors significantly improves the model fit. The difference between the two models holds statistical significance with a p-value of $p < 0.01$, adhering to the predetermined alpha level of 0.05.

5. Discussion

In the following chapter, the key findings of the study are presented, the hypotheses are revisited, and the reliability of the data analysis is discussed. In addition, unexpected results, confounding factors, and biases are discussed and comparisons to previous research in this field are drawn. Moreover, the implications for the future understanding of fragments in German are debated.

5.1 Interpretation of the findings

The results of the study reveal that on the whole, the majority of fragmentary answers were rated as 7 (fully natural). This outcome was somewhat unforeseen and will be subjected to a more comprehensive discussion in this section. Additionally, the current investigation highlighted that factors such as emphasis, modality, and fragment type exerted a noteworthy influence on the perceived naturalness of fragmented responses. These influences will be examined in greater detail in this section as well.

Regarding the fragment type, an unexpected and statistically significant differentiation emerged between functional and lexical fragments. Interestingly, the observed pattern contradicted initial expectations. While it was anticipated that lexical fragments would yield higher acceptability ratings, the opposite proved true: functional fragments received higher ratings. This could be explained by the specific selection of fragment types in the study. Functional fragments were represented by prepositions, while lexical fragments were exemplified by nouns denoting to human referents. It is plausible to consider that a different choice of words might lead to divergent ratings. Nevertheless, given the restricted number of functional words in the German language, particularly those amenable to contrastive analysis,

49

prepositions appear to be the most appropriate candidates. Similarly, nouns signifying human entities are likely the most prevalent type of fragmented responses in everyday discourse, thus representing a fitting benchmark for comparison with functional fragmentary answers.

The lower AIC and significant likelihood ratio test suggest that the more complex model incorporating predictors such as emphasis, modality, and fragment type, provides a better fit to the data. Furthermore, these predictors exert a notable impact on the model's performance. Despite the statistically significant distinction between the AIC of the null model and the model encompassing the three predictors, this discrepancy remains relatively minor. This could potentially be attributed to an unaccounted-for effect within the model or the possibility that modality and emphasis, as predictors, possess relatively weak levels of significance.

From a pilot study that assessed the influence of sociolinguistic factors on how fragmentary answers are perceived by native Dutch speakers, it was discovered that sex and educational background play no significant role, whereas age and geography must be controlled for (cf. Delbar 2019). However, as participants for the present study were crowdsourced from Prolific which has a limited number of potential participants, a prescreening for age and geography was not possible, as the pool of participants would not have been sufficient then. Nevertheless, the data shows that the differences in age are negligible.

Figure 8 shows participants' ratings by age group. Ratings of participants under the age of are coloured in red, while those of participants aged between 30 and 49 are shown in green. Moreover, the ratings of those aged 50 years or older are illustrated in blue and those of participants that did not provide their age are coloured in purple.
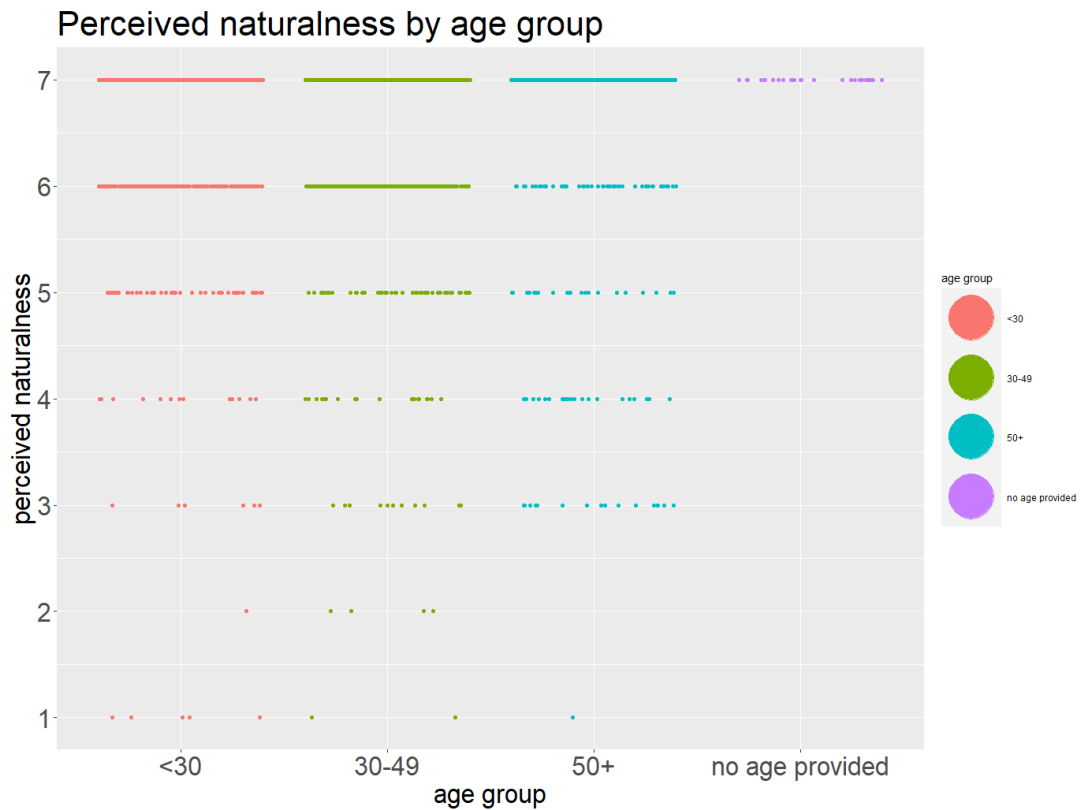
Figure 8: Scatter plot of participants' ratings by age group

As it becomes apparent in Figure 8, there are only slight differences between younger and older participants. Much like the findings of Delbar (2019), only minor age-related differences can be observed. Overall, there seems to be a subtle tendency for participants to give lower acceptability ratings as age increases. That is, participants under the age of 30 provided an average rating of m = 6.66 with a standard deviation of sd = 0.76. Meanwhile, participants aged between 30 and 49 assigned slightly lower average ratings of m = 6.54 with a standard deviation of 0.83. Similarly, participants aged 50 years or older gave the lowest average rating of m = 6.43 with a standard deviation of sd = 1.11. Nonetheless, the data also indicates a noticeable increase in variability within each age group, as age advances. This, in turn, adds a layer of complexity to making comparisons between these groups. Furthermore, the uneven distribution of participants across age groups further compounds the complexity of intergroup comparisons. The age breakdown is as follows: There are 42 participants in the group aged below 30, 41 participants aged between 30 and 49, and only 16 participants aged 50 or older.

While the primary focus of the present study centres on German native speakers, it is noteworthy that a subset of participants indicated proficiency in languages beyond German. When contrasting the acceptability ratings between monolingual and bilingual German speakers, a notable pattern emerges. That is, overall, bilingual participants tend to provide lower ratings. Specifically, bilingual participants yielded an average rating of m = 6.43 with a standard deviation of sd =1.29, whereas monolinguals exhibited slightly higher ratings at m = 6.59 with a standard deviation of sd = 0.80, as is illustrated in Figure 9.
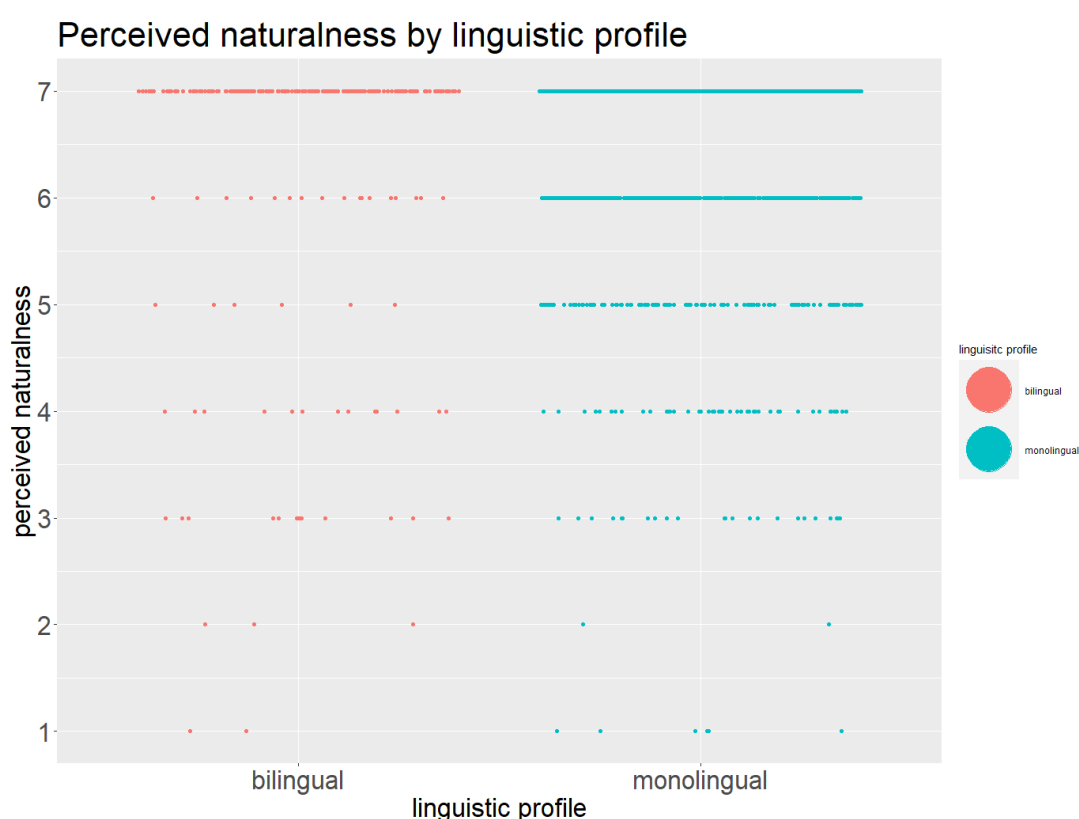
Figure 9: Scatter plot of participants' ratings by linguistic profile

However, it is crucial to exercise caution when drawing broad conclusions from these findings. The higher variability observed in the ratings provided by bilinguals, coupled with their smaller representation in the study necessitates careful consideration. That is, the data set includes only 9 bilinguals compared to the significantly larger group of 91 monolinguals. Further investigation is warranted to explore whether bilingualism truly impacts the perception of fragmentary answers involving contrastive focus.

In terms of regional variations in the acceptance of fragments, the findings of this study should be interpreted within the context of the study's limitations, which restrict the ability to draw comparisons across different regions. The limited number of potential participants available on Prolific has hindered the exploration of potential regional differences. Additionally, the specific setup and methodology employed in this study were designed to assess the effectiveness of conducting acceptability judgement tasks, with a specific focus on factors such as modality, emphasis, and fragment type and have not been configured for capturing nuanced regional variations.

Addressing biases inherent in the study's methodology, it is crucial to recognise several factors that might have influenced participants' responses. Firstly, participants understood their task involved assessing language structures based on their perceived naturalness. While this awareness aligns with the study's objectives, such explicit evaluation might lead to a heightened sensitivity, potentially affecting the participants' judgements.

Furthermore, it is worth noting that participating in a research study inherently deviates from the spontaneous communication that characterises everyday interactions. The conditions inherent to the research setting may potentially trigger different reactions and alter participants' perception of such linguistic phenomena. This alternation might introduce a level of artificiality that contrasts with natural language use. However, the prominence of ratings assigned with a value of 7, indicative of complete naturalness, lends credence to the inference that, despite the controlled framework of the study, participants did not perceive the provided dialogues as inherently unnatural.

This high frequency of ratings indicating a complete naturalness could be attributed to a plausible linguistic phenomenon. That is, the prevalence of occasional linguistic errors in real-life conversations. Native speakers, despite their proficiency, may inadvertently use ungrammatical or unconventional structures in their speech, suggesting that such structures, though not adhering to formal grammar, are still recognisable and accepted as part of natural language use.

Furthermore, it can be said that other factors may come at play when investigating the factors that influence how contrastive, fragmentary answers are perceived by native speakers. First, the difference in emphasis might be clearer if only the contrasting word instead of the complete phrase were given in the fragmentary answer. That is, for functional fragments, the fragmentary answer would only consist of the preposition instead of the prepositional phrase. Similarly, for lexical fragments, the answer would only consist of the contrastive noun instead of the noun phrase.

(28) A: Peter hat AB 18 Uhr im Kino gearbeitet.
            'Peter worked at the cinema FROM 6pm.'
       B: Nein, BIS.
            'No, UNTIL.'

(own example)

(29) A: Peter hat ab 18 Uhr im Kino gearbeitet.
            'Peter worked at the cinema from 6pm.'
       B: Nein, bis.
            'No, until.'

(own example)

As illustrated in (28-29), speaker B's answers can be reduced in such a way that they only incorporate the contrastive words. The dialogues in (28) and (29) differ in whether the contrasting words are emphasised or not. The present study has already shown that contrastive dialogues involving fragmentary answers receive lower ratings in terms of naturalness if the contrasting words are not either orthographically or prosodically emphasised. In such cases, the reader or hearer struggles more to understand the dialogue (29) because they are not primed for the contrast, and it is difficult to understand what speaker B tries to convey with their message. The hearer's task encompasses not only deducing the reference of the preposition *until* from the preceding phrase *until 6pm* but also identifying any contrasts present and subsequently modifying the phrase accordingly. For instance, the hearer needs to recognise and adapt to the contrast in meaning, i.e. *Peter worked at the cinema until 6pm* instead of *Peter worked at the cinema from 6pm.*

In cases such as (28), in which the contrastive words are emphasised, one can assume that is much easier for the hearer to understand what the

preposition refers to, what it contrasts with and what the propositional content of speaker B's utterance is. Therefore, it can be assumed that fragmentary answers with a complete prepositional or noun phrase and emphasised contrastive words receive the highest acceptability ratings. Acceptability ratings lower if the contrastive words lack emphasis. Fragmentary answers without complete phrases and without emphasis receive the lowest acceptability ratings. However, further studies are required to check those hypotheses.

Next, the context might also influence how contrastive, fragmentary answers are perceived. That is, the prominence of the correlate in the discourse possibly impacts how natural the fragmentary answer sound to them. Since no previous context was given when any of the dialogues were presented, it can be ruled out that this affected participants' judgements. Furthermore, all preceding sentences in the critical stimuli had the same syntactic structure and did not include any subordinate clauses. Therefore, it was ensured that all correlates in the preceding sentence had the same degree of prominence in the discourse.

5.2 Comparison with previous studies and theoretical predictions

In this section, the similarities and disparities that emerge between previous research and theoretical conjectures on the one hand, and the empirical outcomes of the current investigation on the other hand are explored.

First, in accordance with the hypotheses drawn from previous research findings outlined in chapter 2, the role of modality in identifying correlate-remnant pairings is apparent, evidenced by the higher acceptability ratings assigned to auditory stimuli. Notably, the auditory stimuli encompassed recordings both with and without emphasis on the contrasting words, necessitating a consideration of the role of emphasis. For an in-depth exploration of this aspect, refer to the latter part of this subchapter.

Returning to the theme of modality, it is important to acknowledge that the manner, in which the contrasting words were emphasised depended on the modality, thereby potentially impacting the ratings. To elaborate, the natural occurrence of pitch accent on contrasting words, as explained in chapter 2.3, contrasts with the less common practice of emphasising words

through uppercase formatting in written language. This distinction could account for the lower ratings received by both emphasised and non-emphasised stimuli in comparison to their auditory counterparts.

Second, as implied by the CRC, dialogues that lack pitch accent on the contrasting words should be rated as less natural. This theoretical supposition aligns with the empirical results, which reveal that stimuli without emphasis indeed receive lower naturalness ratings, and this difference is statistically significant. Despite the statistical significance, it was believed that the difference between ratings assigned to the stimuli with and without emphasis exhibit a more substantial discrepancy.

One could explain this finding as follows. As has already been mentioned in chapter 3.2, it is established that pitch accent inherently occupies the default position, even in cases of silent reading (cf. Rasekhi & Harris, 2021). Consequently, it stands to reason that the same principle applies to pitch accent on contrasting words. Certainly, the inferred pitch accents present during silent reading cannot be equated with those produced in spoken language. Therefore, it is reasonable to conclude that if intonational patterns are somewhat inherent in written language. Hence, it is logical to presume that disparities between modalities exist, although they might not be as substantial as previously believed.

Furthermore, the comparison of auditory and written stimuli underscores the nuanced interaction between emphasis and contrastive focus. Auditory recordings without emphasis still encompass natural intonation, albeit without explicit emphasis on contrasting words. The brevity of the stimuli potentially aids participants in comprehending the underlying contrastive focus even in the absence of explicit emphasis.

Regarding parallelism, it is worth noting that the stimuli exhibited contrasting words that shared both morphological and phonetic characteristics. These encompassed morphological features such as number, gender, case, as well as their prosodic weight. Drawing from the hypothesis on Morphological Parallelism discussed in chapter 2.3, these shared properties contribute to the ease of identifying correlate-remnant pairings. For instance, the case marking ensured the unambiguity of all dialogues, while the similar syntactic structures of the preceding clause and the fragmentary

answer further enhanced the comprehensibility of the dialogues. This parallelism could potentially explain the prevalence of high acceptability ratings observed across the dialogues. However, in order to truly validate the morphological parallelism hypothesis, it is imperative to conduct further research that involves a direct comparison between the outcomes of the current study and dialogues featuring dissimilar correlate-remnant pairings.

Delving into the locality aspect, the design of stimuli intentionally challenges the anticipated correlation between position and accent. By positioning the appropriate correlate away from the final position, which traditionally carries the nuclear sentence accent, the stimuli deviate from the expected norm. Remarkably, despite this intentional violation, acceptability ratings remain consistently high and exhibit minimal variability across conditions. The unexpectedness of these ratings may arise from a complex interplay of linguistic elements and cognitive processing.

Regarding emphasis, as hypothesised, stimuli featuring emphasised contrasting words received higher ratings than those without such emphasis. This trend possibly arises from participants still perceiving a degree of emphasis even in the condition without pitch accent on the correlate. In the written stimuli, on the one hand, participants might have construed it as emphasised due to their mental visualisation of a situation and how the dialogue would unfold in such a scenario, although the dialogue lacked explicit emphasis on the contrasting words. In the auditory stimuli on the other hand, the recordings without emphasis maintained a natural intonation. Here, the emphasis was not directed at the contrasting words but rather at the preverbal position, as exemplified in chapter 3.3. It is plausible that the stimuli sentences were short enough to be understood. That is, participants could understand what the contrastive focus relates to in the preceding clause, without needing the emphasis on the contrasting words. This could provide an explanation for why the ratings of the stimuli that lack emphasis exhibit a statistically significant decrease compared to those featuring emphasis. However, even with this decrease, the ratings for the non-emphasised stimuli do not show a significant drop when compared to the ratings of the emphasised stimuli.

Conducting a comparison of the time participants took to provide their ratings could offer insights into whether understanding the dialogues without emphasis required more time compared to those with emphasis. Extended reading durations might imply that the absence of orthographic marking on the contrasting words makes it comparatively more challenging to establish correlate-remnant relations. However, interpreting the outcomes of studies that rely on reaction times presents a challenge. Longer reaction times can arise from various factors, making the interpretation of cognitive processes less straightforward (cf. Schütze, 2016).

As a last point, the study delved into the potential influence of semantics on the perception of fragmentary answers arising from contrasting words. This investigation aimed to discern whether distinctions between lexical and functional fragments impact their perceived naturalness among native speakers. As explained in chapter 2.3, it is commonly assumed that lexical terms hold a more significant role in conveying meaning and are more likely to carry stress in speech compared to functional words. However, surprisingly, the findings of the current study challenge this convention. The results suggest that contrastive, functional fragments are perceived as more natural in comparison to their lexical counterparts. This outcome could be attributed to the study design, in which prepositions were chosen to represent functional words. In most of the dialogues, those prepositions informed the hearer or reader about specific times or locations. As the dialogues lacked contextual cues, participants might have inferred that details regarding time or location were more important in the conversation than the information given by the lexical words. Therefore, the functional words seemed more prominent in the discourse than the lexical contrasting words. Nevertheless, it is crucial to note that this interpretation remains speculative, necessitating further research to corroborate these findings and delve deeper into the underlying dynamics.

6. Conclusions

The final chapter provides an overview of the findings from the present study, accompanied by an examination of its contributions to the realm of

fragments. In addition, the study's limitations are discussed and an outlook to future research is given.

6.1 Summary of findings

The findings of the study revealed that contrastive fragments are perceived as more natural when the contrasting words are emphasised through prosody or orthography. Furthermore, presenting dialogues as auditory recordings, rather than written text and including functional contrastive fragments in responses as opposed to lexical ones, enhanced the naturalness of contrastive fragments.

The influence of emphasis and modality on participants' ratings was in line with expectations, although the impact was somewhat less pronounced than initially predicted. Regarding emphasis, the present study showed that dialogues in which the contrasting words are either orthographically or prosodically emphasised are perceived as more natural than if they are not. This suggests that emphasis contributes positively to the recipient's ability to establish a focus-based anaphoric relation.

Furthermore, modality is a significant factor for how fragmentary answers are perceived. That is, fragmentary answers receive higher acceptability ratings if presented in auditory format as opposed to their written counterparts. This trend can be attributed to the greater frequency of elliptical utterances in spoken language compared to written language, and the tendency for language structures to appear more natural within their typical contexts.

As hypothesised, the difference between lexical and functional fragments had a significant impact on participants' ratings. However, unexpectedly, it is the dialogues incorporating functional fragments that are rated as more natural than those containing lexical fragments.

Moreover, the results suggest that bilingualism might also impact acceptability ratings, insofar as bilingual speakers generally tend to give lower ratings compared to monolingual speakers.

These findings yield valuable insights into the optimal configuration of stimuli for acceptability judgements, aiming to mitigate the impact of con-

founding variables that could potentially affect the ratings. In summary, acceptability judgement tasks that require participants to identify correlate-remnant pairings should utilise auditory stimuli with emphasis on the correlate and remnant. When including fragments in the stimuli, functional words are recommended, as participants seem to comprehend such dialogues more easily. This approach safeguards that sentences are perceived as natural. Not adhering to this approach might lead participants to find it more challenging to pinpoint the correlate-remnant pairing, resulting in lower ratings for the dialogues. For research questions aiming to identify acceptable structures in a specific language, following the recommended approach helps minimise confounding factors and is therefore advisable.

6.2 Contributions to the field

The present study delves into the factors affecting how native speakers perceive fragments in dialogues featuring contrastive focus in German. This pioneering investigation represents the first empirical examination of hypotheses concerning the placement of pitch accents in dialogues with contrastive focus, setting the stage for future research employing acceptability judgement tasks in this domain. It not only contributes to the existing body of cross-linguistic research by exploring contrasting fragments in German but also enhances the collective knowledge by validating earlier hypotheses and research discoveries from other languages in the context of German.

As discussed in chapter 2.4, previous instances of acceptability judgement tasks have often relied on informal approaches. Consequently, a need arises for formally gathered data through acceptability judgement tasks that meet the criteria essential for scientific research. This study addresses this demand by recruiting 100 participants, ensuring robust statistical power, and employing a 7-point Likert scale to encompass an optimal array of response options. In doing so, this study diligently aligns with the prerequisites established for conducting structured and formal acceptability judgement tasks, as outlined in chapter 2.4.

Moreover, the study contributes to the existing research on the significance of orthographic marking. While this approach is already established

and utilised in other linguistic fields, this study stands out as the first to investigate the role of orthography in comprehending contrastive focus.

In addition, the study introduces new insights into the processing of functional fragments. Results indicate that functional words play a more pivotal role than previously believed. While lexical words were traditionally considered vital for conveying messages, as outlined in chapter 2.3, this study highlights the prominence of functional fragments. This is evidenced by higher acceptability ratings for dialogues containing functional fragments compared to those with lexical fragments.

Furthermore, as explained in chapter 2.3, prosodic emphasis is inherently more natural than the utilisation of orthographic markings, and the linguistic phenomenon of elliptical structures finds greater prevalence in spoken language as opposed to written language. The results of the study indicate that auditory stimuli prove to be a more suitable choice for conducting acceptability judgement tasks within this specific field. While conducting acceptability judgement tasks with auditory recordings demands more resources, the study underscores the necessity of this approach and its capacity to yield more reliable results in linguistics.

Finally, the results partially support the hypothesis regarding clause parallelism. Elliptical structures with morphological and syntactic similarity to the preceding sentence garnered higher acceptability ratings. However, it is important to note that dissimilar structures were not investigated in this study, and thus, the parallelism hypothesis could not be fully tested.

6.3 Limitations of the study

In the following, any potential limitations or biases that might have impacted the research are discussed. First, in this study, participant recruitment was conducted through the Internet. This introduces certain biases, insofar this approach entails that only individuals with Internet access, available time for survey completion, and a willingness to participate are included. Therefore, the recruited participants might not be perfectly representative of all German native speakers. Nevertheless, Sprouse (2011) and others have shown that conducting AJTs through crowdsourcing platforms such as Amazon's Me-

chanical Turk are almost indistinguishable from the laboratory data. However, since Amazon's Mechanical Turk was not well-suited for the present study due to certain setup constraints, such as the requirement for a US credit card and US social security number, Prolific emerges as a viable alternative (cf. Häussler & Juzek, 2016).

Moreover, the study's scope was limited to the German language. While the findings should not be broadly generalised to contrastive fragments in other languages, they do reveal indicative patterns that can serve as a foundation for formulating hypotheses within language contexts related to German. Nonetheless, the validation of these hypotheses necessitates additional cross-linguistic investigations.

On the one hand, the present study exclusively examined contrastive fragments, rendering its findings applicable primarily to this specific category of correlate-remnant pairings. On the other hand, the research concentrated on a restricted selection of functional words, specifically investigating only three distinct sets of contrasting preposition pairs. This narrow focus highlights that the extent of exploration on functional words is relatively limited within this study.

Moreover, it cannot be ruled out that there might be potential confounding factors that may have influenced the results and were not addressed within the statistical analysis. For instance, age-based differences in acceptability ratings or variations in regional dialects could potentially introduce confounding influences. While neither age nor dialect variations were the central focus of the present study, their emergence as potential factors necessitates future research to delve into these aspects.

In conclusion, this examination of limitations and biases underscores the need for a more comprehensive understanding and subsequent research in order to gain a clearer perspective on the broader implications of the findings.

6.4 Suggestions for future research

The following section delves into aspects of the study that warrant further investigation. First, the results did not align with the hypothesis that lexical words hold greater prominence in discourse, thereby making contrastive

lexical fragments easier to process. Instead, functional fragments received higher acceptability ratings, somewhat challenging prior research. This discrepancy highlights the necessity for future studies that delve into the contrasts between lexical and functional words in fragmentary responses.

Second, as predicted, dialogues featuring both the correlate and remnant bearing pitch accents were rated as more natural compared to cases where only the correlate lacked prosodic emphasis. However, the latter dialogues still garnered relatively high acceptability ratings. Consequently, future research should undertake a more comprehensive examination to determine if emphasising both contrasting words is necessary or if solely applying pitch accent to the remnant, as outlined in the CRC, for identifying correlate-remnant pairings is sufficient.

Moreover, the results suggest a potential influence of bilingualism on acceptability judgement ratings. Further exploration is warranted to ascertain whether possessing multiple native languages impacts the perception of natural language structures.

In addition, the investigation into how different modalities affect the perception of utterances remains incomplete without the inclusion of sign language. Given that the current study exclusively considers written and auditory modalities, it is imperative to conduct additional research that compares the current findings with studies involving signed utterances.

Further investigations should also explore the significance of having a complete prepositional phrase in the fragmentary answer. They should focus on dialogues that present fragmentary responses comprising solely the prepositions, allowing for a comparative analysis against the current study's findings. This approach will help address whether native speakers possess the capability to effectively comprehend fragmentary responses that exclusively feature the contrasting preposition. In conclusion, while more research is undoubtedly needed, the present study already yields valuable insights into the comprehension of contrastive fragments within the German language.

7. References

Akinnaso, F. N. (1982). On The Differences Between Spoken and Written Language. *Language and Speech*, *25*(2), 97–125. https://doi.org/10.1177/002383098202500201

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Beckman, M. E., & Ayers. (1997). *Guidelines for ToBI labelling, version 3*. The Ohio State University Research Foundation. http://www.ling.ohio-state.edu/phonetics/ToBI/ToBI.0.html

Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). *Grammar of Spoken and Written English*. John Benjamins Publishing Company. https://doi.org/10.1075/z.232

Black, M., Coltheart, M., & Byng, S. (1985). Forms of coding in sentence comprehension during reading. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 655–672). Lawrence Erlbaum Associates.

Carlson, K., Frazier, L., & Clifton, C. (2009). How prosody constrains comprehension: A limited effect of prosodic packaging. *Lingua, 119*(7), 1066–1082. https://doi.org/10.1016/j.lingua.2008.11.003

Chomsky, N. (1957). *Syntactic Structures*. Mouton de Gruyter.

Chomsky, N., & Halle, M. (1991). *The sound pattern of English* (1st MIT Press paperback. ed). MIT Press.

Christensen, R. H. B. (2018). *Regression Models for Ordinal Data: Intro-ducing R-package ordinal* [Computer software]. https://cran.r-pro-ject.org/package=ordinal

Curtis, A., Smith, T., Ziganshin, B., & Elefteriades, J. (2016). The Mystery of the Z-Score. *AORTA*, *4*(4), 124–130. https://doi.org/10.12945/j.aorta.2016.16.014

Delbar, N. A. (2019). *Swiping in English and Dutch: The Interaction be-tween R-Pronouns and Modal Particles*.

Featherston, S. (2008). Thermometer judgements as linguistic evidence. In C. M. Riehl & A. Rothe (Eds.), *Was ist linguistische Evidenz?* Shaker Verlag.

Féry, C. (2011). German sentence accents and embedded prosodic phrases. *Lingua*, *121*(13), 1906–1922. https://doi.org/10.1016/j.lin-gua.2011.07.005

Féry, C., & Herbst, L. (2004). German Sentence Accent Revisited. *Inter-disciplinary Studies in Information Structures 1. Working Pa-Pers of the SFB 632*, 43–75.

Franck, J., Bowers, J., Frauenfelder, U. H., & Vigliocco, G. (2003). Ortho-graphic influences on agreement: A case for modality-specific form effects on grammatical encoding. *Language and Cognitive Pro-cesses*, *18*(1), 61–79. https://doi.org/10.1080/01690960143000452

Frazier, L., & Clifton, C. (1998). Comprehension of Sluiced Sentences. *Language and Cognitive Processes*, *13*(4), 499–520. https://doi.org/10.1080/016909698386474

Frazier, L., Taft, L., Roeper, T., Clifton, C., & Ehrlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition*, *12*(5), 421–430. https://doi.org/10.3758/BF03198303

Griffiths, J., Güneş, G., & Lipták, A. (2023). Reprise fragments in English and Hungarian: Further support for an in-situ Q-equivalence approach to clausal ellipsis. *Language*, *99*(1), 154–191. https://doi.org/10.1353/lan.2023.0000

Griffiths, J., & Lipták, A. (2014). Contrast and Island Sensitivity in Clausal Ellipsis. *Syntax*, *17*(3), 189–234. https://doi.org/10.1111/synt.12018

Harris, J. A. (2015). Structure Modulates Similarity-Based Interference in Sluicing: An Eye Tracking study. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01839

Harris, J. A., & Carlson, K. (2016). Keep it local (and final): Remnant preferences in "let alone" ellipsis. *Quarterly Journal of Experimental Psychology*, *69*(7), 1278–1301. https://doi.org/10.1080/17470218.2015.1062526

Harris, J. A., & Carlson, K. (2018). Information Structure Preferences in Focus-Sensitive Ellipsis: How Defaults Persist. *Language and Speech*, *61*(3), 480–512. https://doi.org/10.1177/0023830917737110

Häussler, J., & Juzek, T. (2016). Hot Topics Surrounding Acceptability Judgement Tasks. *Proceedings of Linguistic Evidence*. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/77638

Hedderich, J., & Sachs, L. (2016). *Angewandte Statistik*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45691-0

Jasso, J. (2022). *How Does Parent Input Influence Bilingual Children's Knowledge and Use of Spanish Subjunctive? A Dyadic Study* [Dissertation]. University of Texas at Austin.

Juzek, T. S. (2016). *Acceptability Judgement Tasks and Grammatical Theory*. University of Oxford.

Kayali, N. (2023). "Does this make sense?": The effect of matching guise in regional accent on grammatical acceptability judgments. *Proceedings of the Linguistic Society of America*, *8*, 5525. https://doi.org/10.3765/plsa.v8i1.5525

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*(3–4), 243–276. https://doi.org/10.1556/ALing.55.2008.3-4.2

Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press.

Lobeck, A. C. (1995). *Ellipsis: Functional heads, licensing, and identification*. Oxford University Press.

Merchant, J. (2004). Fragments and ellipsis. *Linguistics and Philosophy*, *27*(6), 661–738. https://doi.org/10.1007/s10988-005-7378-3

Merchant, J. (2019). Ellipsis: A survey of analytical approaches. In J. Van Craenenbroeck & T. Temmerman (Eds.), *The Oxford Handbook of Ellipsis* (1st ed., pp. 19–45). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198712398.013.2

Meurers, D., Ziai, R., Amaral, L., Boyd, A., Dimitrov, A., Metcalf, V., & Ott, N. (2010). Enhancing Authentic Web Pages for Language Learners. *Proceedings of the 5th Workshop on Innovative Use of NLP for Building Educational Applications, NAACL-HLT 2010*, 10–18. http://purl.org/dm/papers/meurers-ziai-et-al-10.html

Opitz, A., & Bordag, D. (2022). The Impact of Orthography on Lexical Access: The Case of Capitalization and Word Category Information in L1 and L2 German. *Studies in Second Language Acquisition*, *44*(4), 1194–1209. https://doi.org/10.1017/S0272263121000711

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Phillips, C., & Parker, D. (2014). The psycholinguistics of ellipsis. *Lingua*, *151*, 78–95. https://doi.org/10.1016/j.lingua.2013.10.003

Prolific Academic. (2019). [Computer software]. https://prolific.ac/

R Development Core Team. (2015). [Computer software]. http://www.r-project.org/

Rasekhi, V., & Harris, J. A. (2021). Resolving ambiguous polarity stripping ellipsis structures in Persian. *Glossa: A Journal of General Linguistics*, *6*(1), 1–31. https://doi.org/10.16995/glossa.5881

Schütze, C. T. (2016). The empirical base of linguistics: Grammaticality judgments and linguistic methodology [Application/pdf]. *Classics in Linguistics*, 1.01 MB. https://doi.org/10.17169/LANGSCI.B89.100

Sedarous, Y., & Namboodiripad, S. (2020). Using audio stimuli in accepta-

    bility judgment experiments. *Language and Linguistics Compass*,

    14(8), 1–21. https://doi.org/10.1111/lnc3.12377

Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collec-

    tion of acceptability judgments in linguistic theory. *Behavior Re-*

    search Methods, 43(1), 155–167. https://doi.org/10.3758/s13428-

    010-0039-7

Sprouse, J., Schütze, C. T., & Almeida, D. (2013). A comparison of infor-

    mal and formal acceptability judgments using a random sample

    from Linguistic Inquiry 2001–2010. *Lingua, 134*, 219–248.

    https://doi.org/10.1016/j.lingua.2013.07.002

Ståhle, L., & Wold, S. (1989). Analysis of variance. *Chemometrics and In-*

    telligent Laboratory Systems, *6*(4), 259–272.

    https://doi.org/10.1016/0169-7439(89)80095-4

Wagner, M. (2012). Focus and givenness: A unified approach. In I.

    Kučerová & A. Neeleman (Eds.), *Contrasts and Positions in Infor-*

    mation Structure (1st ed., pp. 102–147). Cambridge University

    Press. https://doi.org/10.1017/CBO9780511740084.007

Wagner, P. S. (1999). The synthesis of German contrastive focus. *Pro-*

    ceedings of the 14th ICPhS, 1529–1532.

Wierzba, M., Brown, J. M. M., & Fanselow, G. (2023). The syntactic flexi-

    bility of German and English idioms: Evidence from acceptability

    rating experiments. *Journal of Linguistics*, 1–38.

    https://doi.org/10.1017/S0022226723000105

Winkler, S. (2019). Ellipsis and Prosody. In J. Van Craenenbroeck & T.

    Temmerman (Eds.), *The Oxford Handbook of Ellipsis* (1st ed., pp.

    357–386). Oxford University Press. https://doi.org/10.1093/ox-

    fordhb/9780198712398.013.15

8. Appendix

8.1 Abbreviations, symbols, and other notational conventions

?	questionable/marginal acceptability
*	ungrammatical
//	intonational phrase boundary
UPPERCASE	emphasis (orthographical marking or pitch accent)
[...]$_F$	focused position
α	significance level
β$_1$	coefficient associated with the predictor
A, B	speaker
AIC	Akaike Information Criterion
AJT	acceptability judgement task
ANOVA	analysis of variance
CLMM	Cumulative Link Mixed Models
CRC	Contrastive Remnant Condition
DP	determiner phrase
H*	high pitch
L	low tone
L+H*	pitch accent
m	mean
sd	standard deviation

8.2 List of critical items

The following list of critical items only includes written items. The auditory critical items can be found on this page: https://shorturl.at/JLSW9.

(1)   A: Peter hat seinem BRUDER ein Buch geschenkt.

            B: Nein, seinem VATER.

(2)   A: Peter hat seinem Bruder ein Buch geschenkt.

            B: Nein, seinem Vater.

(3)   A: Peter hat dem POLIZISTEN seinen Ausweis gezeigt.

            B: Nein, dem TÜRSTEHER.

(4)   A: Peter hat dem Polizisten seinen Ausweis gezeigt.

            B: Nein, dem Türsteher.

(5)   A: Peter hat seinem CHEF den neuen Mitarbeiter vorgestellt.

            B: Nein, seinem KOLLEGEN.

(6)   A: Peter hat seinem Chef den neuen Mitarbeiter vorgestellt.

            B: Nein, seinem Kollegen.

(7)   A: Peter hat dem MALER ein Getränk angeboten.

            B: Nein, dem GÄRTNER.

(8)   A: Peter hat dem Maler ein Getränk angeboten.

            B: Nein, dem Gärtner.

(9)   A: Peter hat seinem KOLLEGEN Urlaubsbilder gezeigt.

            B: Nein, seinem NACHBARN.

(10)   A: Peter hat seinem Kollegen Urlaubsbilder gezeigt.

            B: Nein, seinem Nachbarn.

(11)   A: Peter hat seinem NEFFEN Werkzeug geschenkt.

            B: Nein, seinem NACHBARN.

(12)   A: Peter hat seinem Neffen Werkzeug geschenkt.

            B: Nein, seinem Nachbarn.

(13)   A: Peter hat seinem VORGESETZTEN einen Kaffee gebracht.

            B: Nein, seinem MITBEWOHNER.

(14)   A: Peter hat seinem Vorgesetzten einen Kaffee gebracht.

            B: Nein, seinem Mitbewohner.

(15)   A: Peter hat AB 18 Uhr im Kino gearbeitet.

            B: Nein, BIS 18 Uhr.

(16)   A: Peter hat ab 18 Uhr im Kino gearbeitet.

B: Nein, bis 18 Uhr.

(17)   A: Peter hat BIS August Miete gezahlt.

        B: Nein, AB August.

(18)   A: Peter hat bis August Miete gezahlt.

        B: Nein, ab August.

(19)   A: Peter hat MIT seinem Bruder Unterschriften gesammelt.

        B: Nein, OHNE seinen Bruder.

(20)   A: Peter hat mit seinem Bruder Unterschriften gesammelt.

        B: Nein, ohne seinen Bruder.

(21)   A: Peter hat OHNE sein Team einen Vortrag gehalten.

        B: Nein, MIT seinem Team.

(22)   A: Peter hat ohne sein Team einen Vortrag gehalten.

        B: Nein, mit seinem Team.

(23)   A: Peter hat VOR seiner Mittagspause seine Chefin angerufen.

        B: Nein, NACH seiner Mittagspause.

(24)   A: Peter hat vor seiner Mittagspause seine Chefin angerufen.

        B: Nein, nach seiner Mittagspause.

(25)   A: Peter hat NACH seinem Urlaub den Handwerker gerufen.

        B: Nein, VOR seinem Urlaub.

(26)   A: Peter hat nach seinem Urlaub den Handwerker gerufen.

        B: Nein, vor seinem Urlaub.

(27)   A: Peter ist VOR seinem Einkauf noch zur Bank gegangen.

        B: Nein, NACH seinem Einkauf.

(28)   A: Peter ist vor seinem Einkauf noch zur Bank gegangen.

        B: Nein, nach seinem Einkauf.

8.3 List of filler items

The following list of filler items only includes written items. The auditory filler items can be found on this page: https://shorturl.at/drFO0. The acceptability of the fillers varied, with A representing full acceptability, B indicating some acceptability, C denoting neutrality in terms of acceptability, D implying partial unacceptability, and E signifying complete unacceptability.

A1   A: Peter hat in der Mensa zu Mittag gegessen.

        B: Ja, zusammen mit Freunden.

A2   A: Peter hat den Gegenspieler vorsätzlich gefoult.

      B: Ja, den Stürmer.

A3   A: Peter hat die SÜDDEUTSCHE gelesen.

      B: Nein, er hat die FAZ gelesen.

A4   A: Peter hat einen ERDBEERKUCHEN gebacken.

      B: Nein, er hat einen SCHOKOKUCHEN gebacken.

A5   A: Peter hat den KAFFEE gekocht.

      B: Nein, er hat den TEE gekocht.

B1   A: Peter hat dem Fürsten jemanden empfohlen.

      B: Ja, dem Fürsten den Maler.

B2   A: Peter hat dem Gast ein Getränk empfohlen.

      B: Ja, dem Gast den Wein.

B3   A: Peter hat seinem Neffen ein Geschenk gegeben.

      B: Ja, seinem Neffen ein Fahrrad.

B4   A: Peter hat geglaubt, dass sein CHEF Urlaub hat.

      B: Nein, er hat geglaubt, sein Chef gibt IHM Urlaub.

B5   A: Peter hat sich GEWUNDERT, weil Maria zu Besuch kam.

      B: Nein, er hat sich GEFREUT, weil Maria hat Geschenke
        mitgebracht.

B6   A: Peter hat angenommen, dass Franz ihm das Radio
        SCHENKT.

      B: Nein, er hat angenommen, er VERKAUFT ihm das Radio
        günstiger.

C1   A: Peter hat dem Kunden etwas gezeigt.

      B: Ja, dem Kunden sich selbst im Spiegel.

C2   A: Peter hat den Mann nach etwas gefragt.

      B: Ja, wen wer in dieser Affäre betrügt.

C3   A: Peter hat seinen Nachbar zu dem Unfall befragt.

      B: Ja, wem wer aufgefahren ist.

C4   A: Peter hat gedacht, dass der POLITIKER bestochen wurde.

      B: Nein, in ROTTENBURG hat Peter gedacht, hat der Händler
        den Politiker bestochen.

C5   A: Peter hat erzählt, dass Franz einen UNFALL hatte.

      B: Nein, auf einer KREUZUNG hat Peter erzählt, hatte Franz

einen Unfall.

C6   A: Peter hat gehört, dass der Lehrer WÄHREND seinem Urlaub gekündigt hat.

B: Nein, VOR dem Urlaub hat Peter gehört, hat der Lehrer ge kündigt.

D1   A: Peter hat ihn als kompetenten Begleiter empfohlen.

B: Ja, sich selbst.

D2   A: Peter hat Maria einen Brief geschrieben.

B: Ja, einander.

D3   A: PETER hat es dem neuen Tenor zugemutet.

B: Nein, der KOMPONIST hat dem neuen Tenor es zugemutet.

D4   A: Peter hat seinen Sohn eine GESCHICHTE vorgelesen.

B: Nein, Peter hat ein GEDICHT ihm vorgelesen.

D5   A: Peter hat Maria eine E-MAIL geschickt.

B: Nein, er hat eine SMS ihr geschickt.

D6   A: Peter hat am liebsten die FAZ gelesen.

B: Nein, er liest am liebsten die SÜDDEUTSCHE, obwohl er lebt jetzt in Düsseldorf.

E1   A: Peter hat den Rasen gemäht.

B: Ja, obwohl der Hitze.

E2   A: Peter hat den Fernseher eingeschaltet.

B: Ja, um zu sehen eine Fernsehserie.

E3   A: Peter hat seinem Sohn ein Geschenk gemacht

B: Ja, ein Fahrrad in die Schule zum Fahren.

E4   A: Peter hat mit Freunden UNO gespielt.

B: Nein, beim Stammtisch die Freunde haben mit Vorliebe SKAT gespielt.

E5   A: Peter hat Franz mit einem Geschenk überrascht.

B: Nein, da gerechnet mit hat der Franz natürlich nicht.

Declaration of Authorship

I hereby confirm that this paper and the work presented in it is entirely my own. Where I have consulted the work of others this is always clearly stated. All statements taken literally from other writings or referred to by analogy are marked and the source is always given. This paper has not yet been submitted to another examination office, either in the same or similar form.

Tübingen, September 20th, 2023

_____

Miriam Schiele