COMP-trace effects in German: the role of processing

Ankelien Schippers[1], Margreet Vogelzang[2] & David Öwerdieck[3]
[1] *University of Oldenburg,* [2] *University of Cambridge,* [3] *TU Braunschweig*

Abstract

This article reports on the processing and comprehension of COMP-trace violations in German. The status of the COMP-trace effect in German is a controversial issue. It has been argued that judgments on long-distance (LD) subject questions are distorted because of parsing problems in the main clause, the embedded clause, or both, and that LD subject questions are sometimes misinterpreted as object questions. Our self-paced reading data shows that processing difficulties with LD subject questions occur in the embedded clause, not the main clause, particularly at the point at which an embedded subject gap is postulated. Our study furthermore shows that readers erroneously interpret subject as object LD questions, but only when the embedded clause contains a case-ambiguous DP. A case-ambiguous DP thus functions as a superficial work-around for a COMP-trace violation. As we argue, our data support the view that German has a genuine COMP-trace effect and that potential parsing problems only occur in the context of local ambiguities. We propose that differences in the magnitude and fatality of COMP-trace violations between languages can be explained by formulating the COMP-trace effect in terms of accessibility, rather than a categorical syntactic constraint.

1. Introduction[1]

In many languages, subjects are harder to extract from embedded clauses than objects. In English, this emerges as the so-called COMP-trace effect, which refers to the fact that embedded subject traces cannot be immediately preceded by a complementizer. Long-distance (LD) subject questions as in (1) differ from LD object questions as in (2) in this respect, where the complementizer is optional:

(1)   Which author do you think (*that) __ appreciated the publisher?
(2)   Which author do you think (that) the publisher appreciated __?

Other languages require special complementizers for subject questions, such as French (Perlmutter 1971) and West Flemish (Haegeman 1983), or resumptive pronouns (e.g. Hebrew; Rizzi & Shlonsky 2007). Finally, there are also languages that solve the problem by moving the entire embedded CP (clausal pied piping, attested in Imbabura Quechua; Cole & Hermon 1981; Hermon 1984). The consensus in syntactic theory therefore appears to be that subject positions are 'special' in some sense and cannot be vacated without problems, and that whatever constraint is responsible for this holds cross-linguistically (Rizzi & Shlonsky 2007). However, a stubborn problem in this respect is that some languages do appear to allow COMP-trace violations. German and Dutch are notable examples of such languages, as the German example in (3) and the Dutch example in (4) show. In Dutch, LD subject movement from *dat*-clauses is completely grammatical. In German, it is associated with decreased acceptability, but also not ruled out altogether.

---

(3)   ?Welch-er Schriftsteller meinst du,   dass   den   Verleger geschätzt   hat?
                which-NOM author   think you   that   the.ACC publisher appreciated has?
                'Which author do you think appreciated the publisher?'

(4)   Welke auteur denk je   dat   hem   gewaardeerd heeft?
                Which author think you   that   him   appreciated has?
                'Which author do you think appreciated him?'

In this paper, the focus is on German. We specifically look at how sentences with COMP-trace violations are processed and comprehended by means of a self-paced reading and comprehension task. Our study is novel in this respect, since it is the first to look at this kind of data for these constructions. We test different hypotheses that have been proposed to explain the lower sensitivity to COMP-trace violations in German. In particular, it has been proposed that COMP-trace violations in German do not occur and that the lower acceptability of LD subject extraction is due to a parsing problem in the main clause (Haider 2007). Alternatively, it has been proposed that COMP-trace violations do occur in German, but that judgments on it are fuzzy because subject and object questions are relatively difficult to distinguish since word-order doesn't differentiate between readings (Featherston 2005; Kiziak 2010). What our data show is that LD subject questions in German are associated with more processing difficulty in the embedded clause only, and that temporary case ambiguities in the embedded clause lead to frequent misinterpretations for subject questions. Based on these results, we argue that the COMP-trace effect in German cannot be simply reduced to (temporary) main clause parsing problems, but that the problem is located in the embedded clause, particularly at the point immediately following the complementizer. We propose that the relevant difference between a language like German on the one hand and English on the other is best explained by formulating the problem with LD subject extraction in terms of accessibility, with subjects being less accessible for LD movement than non-subjects. The relevant difference between languages is whether the inaccessibility for embedded subjects is grammaticalized or not. The outline of the paper is as follows. In section 2, we summarize the main syntactic analyses of the COMP-trace effect. In section 3, we present the method of our study. The results are presented in section 4, after which we end with a discussion and conclusions in section 5 and 6, respectively.

2. COMP-trace

A wide variety of explanations have been offered for the COMP-trace effect. Often these analyses primarily focus on English and do not necessarily account for a ban on LD subject extraction in other languages. In addition to syntactic accounts, there are also prosodic (Kandybowicz 2006, 2008; Ha 2010; Sato & Dobashi 2016; McFadden & Sundaresan 2017), informational-structural (Bennis 1986; Bayer 2005; Bayer & Salzmann 2013) and processing (Culicover 1993; Hawkins 1999, 2004) explanations for the COMP-trace effect. Furthermore, McDaniel et al. (2015) have recently given a production-based account of the COMP-trace effect. In this section, we focus on syntactic accounts of the COMP-trace effect, since these will serve as the starting point for our discussion. The earliest discussion of the COMP-trace effect simply described it in terms of a filter, stating that a configuration of a complementizer followed by trace is disallowed (Perlmutter 1971; Bresnan 1972; Chomsky & Lasnik 1977). A more principled syntactic explanation, and for a long time the dominant analysis, is known as the Empty Category Principle (ECP: Chomsky 1981, 1986; Lasnik & Saito 1984, 1992; Rizzi 1990, see also Pesetsky 1982, which can be seen as a precursor to this explanation). The ECP states that traces have to be properly governed, either through antecedent government or head government. Subject traces are not head-governed and therefore require antecedent government. It is assumed that complementizers block antecedent government, explaining why they cannot occur with a subject trace. More recently, the ECP was recast in terms of a freezing effect (Rizzi & Shlonsky 2007), the idea being that there is a Subject Criterion which forces subjects to move to a criterial position, where they are 'frozen' in place, meaning they cannot move further up. In order to create an LD subject dependency, the criterial subject position either has to be skipped, or the Subject Criterion must be fulfilled

in some other way. For English, it is assumed that null complementizers can fulfil the Subject Criterion in ways that overt complementizers cannot. Overt complementizers therefore cause freezing effects, i.e. they prevent the subject from moving further up (other Freezing-type accounts have been proposed by Pesetsky & Torrego 2001 and Boeckx 2008).

Another type of syntactic explanation of the COMP-trace effect is the so-called anti-locality account (Ishii 2004; Brillman & Hirsch 2014; Bošković 2016; Douglas 2017 and many others). Here, it is assumed that the problem with LD subject movement is that it involves a movement step that is 'too local'. Under standard assumptions, subjects are located in the specifier of TP and have to move to the specifier of the immediately dominating CP on their way to their final landing site. The initial movement step from TP to CP is assumed to be derivationally ruled out because it is too short. The idea is that complementizerless clauses have a reduced CP or no CP at all, so that an anti-locality violation is prevented from occurring.

Both freezing accounts and anti-locality accounts (implicitly) view the constraint responsible for COMP-trace as dichotomous, i.e. as inducing absolute ungrammaticality. In cases where it is apparently possible to violate this constraint (as is, for example, the case for certain speakers/'dialects' that apparently do allow COMP-trace violations, cf. Sobin 1987), it is assumed that some sort of special strategy is being employed by which the relevant constraint is sidestepped. A major problem with all syntactic accounts of the COMP-trace effect is that they have difficulty accounting for cases where the COMP-trace effect appears to be gradual, rather than categorical. As we will argue in the next section, this appears to be exactly the case in German: COMP-trace configurations are bad, but not ruled out altogether. Under the current leading syntactic analyses of the COMP-trace effect (i.e. anti-locality and freezing accounts), this is a puzzling state of affairs. Why would an anti-locality violation lead to ungrammaticality in one language (English) but decreased acceptability in another (German)? Similarly, why are embedded subjects frozen in English, but partly 'defrosted' in German?

2.1   COMP-trace in German

The existence of a COMP-trace effect in German has been contested, with some arguing that German does not have a COMP-trace effect at all (Müller & Sabel 1989; Haider 2010). Matters are somewhat complicated since LD movement constructions in German have a degraded status for many speakers, regardless of whether a subject or a non-subject has been moved. The acceptability of LD movement has a diachronic as well as a dialectal dimension (see Schippers 2012 for a discussion). The correct generalization seems to be that, especially for speakers from northern Germany, LD movement is no longer a productive strategy. Still, judgment experiments have shown that regardless of the degraded status of LD movement in general, there is a clear subject/object asymmetry (Featherston 2005; Kiziak 2010) in the sense that LD subject movement over a complementizer is significantly degraded compared to LD object movement. No such asymmetry exists for complementizerless clauses (also called 'embedded V2 clauses'). The relative patterns of acceptability in German and English are therefore strikingly similar, the most important differences being that LD movement in German has a much more degraded status, and that LD subject movement does not lead to downright ungrammaticality in German (Haider 2010). As pointed out at the end of the previous section, this is a puzzling state of affairs: why are COMP-trace violations bad, but not completely ruled out in German? Haider (2007) claims that COMP-trace violations do not arise in German at all. He assumes that subjects are directly extracted from VP and thus do not leave a trace immediately following the complementizer. He argues that the relatively low acceptability of LD subject movement in German is parsing-related. Specifically, he argues that readers attempt to integrate a subject *wh*-phrase with the matrix verb, which causes an agreement clash. This agreement clash is absent for object questions and presumably causes a drop in acceptability for subject LD-questions (see Andersson & Kvam 1984; Hawkins 1999 and Fanselow & Frisch 2006 for similar ideas). Kiziak (2010), on the other hand, does assume German LD subject questions incur COMP-trace violations, but that judgments in German are fuzzier than in English due to word order differences. In English, word order in the embedded clause crucially distinguishes subject from object questions (subject gaps leave an empty position to the left of the verb,

object gaps to the right), but in German, subject and object gaps both occur to the left of the verb, and Germans therefore have to rely on morphosyntactic cues to distinguish subject from object questions (i.e. case-marking and/or subject-verb agreement). Under the assumption that readers do not always form complete and accurate syntactic and semantic representations of sentences ('good enough' processing, cf. Ferreira & Patson 2007), subject questions could therefore sometimes be misinterpreted as object questions. This could effectively reduce the difference in acceptability between LD subject and object questions.[2] Kiziak suggests such misinterpretations are more likely to occur in the context of temporary case ambiguities. In her acceptability judgment experiments, Kiziak also tested LD questions in which either the *wh*-phrase itself was case-ambiguous (5), or the argument in the embedded clause (6) was:

(5) a. Welche Lehrer-in   denkst du, dass den   Schüler angeschrien hat?
         Which   teacher-FEM think you   that   the.ACC student   yelled.at   has
         'Which teacher do you think yelled at the student?'
     b. Welche Lehrer-in   denkst du, dass der   Schüler angeschrien hat?
         Which   teacher-FEM think you   that   the.NOM student   yelled.at   has
         'Which teacher do you think the student yelled at?'

(6) a. Welch-er   Herzog denkst du, dass die Königin belogen hat?
         Which-NOM duke   think you   that   the queen   lied.to   has
         'Which duke do you think lied to the queen?'
     b. Welch-en   Herzog denkst du, dass die Königin belogen hat?
         Which-ACC duke   think you   that   the queen   lied.to has
         'Which duke do you think the queen lied to?'

Her data show that the difference in acceptability between LD subject and object questions becomes smaller when either the *wh*-phrase or the embedded DP is case-ambiguous. Thus, there is indeed evidence that judgments on the COMP-trace effect become fuzzier with temporary case ambiguities, presumably due to parsing effects, but this should be verified by looking at how such questions are processed and comprehended. Our study fills an important gap in this respect. Secondly, the question is whether (temporary) misparses also occur in LD questions with case-unambiguous arguments; in other words, whether the identical word order for LD subject and object questions alone can result in misinterpretations. Again, we need online processing data and comprehension data to answer this question.

Summarizing, there are two parsing-related explanations for the reduced COMP-trace effect in German; one locating the problem in the main clause (Haider), the other in the embedded clause (Kiziak). Under the first hypothesis, German does not have a COMP-trace effect, since it is assumed that subject traces are not adjacent to the complementizer. Under the second hypothesis, German does have a COMP-trace effect, but judgments on it are fuzzier than in English. Neither Haider's nor Kiziak's hypotheses have been tested by looking at how LD subject and object questions in German are processed and comprehended. The current study fills an important gap in this respect, since the status of the COMP-trace effect in German is still under debate. Therefore, we carried out a self-paced reading task in which each LD question was followed by a comprehension task. This experiment is described in full detail in the next section.

---

[2] An anonymous reviewer points out that another factor that could obfuscate the COMP-trace effect in German is the possibility to scramble objects over subjects in German, so that the subject trace is in a lower position. Kiziak (2010) tested this experimentally but found no supporting evidence for an ameliorating effect of scrambling, though she remarks that the effect could simply be too weak to be picked up in the experiment. Alternatively, it could be the case that the ameliorating effect of scrambling on the COMP-trace effect is offset by the relative markedness of object scrambling.

3. The current study

We address the following research questions with our experiment:
  (1) Are German LD subject questions misinterpreted as object questions?
  (2) Do such misinterpretations occur more frequently in the context of local case ambiguities?
  (3) Are German LD subject questions more difficult to process than object questions and is the processing difficulty located in the main clause, the embedded clause, or both?

With respect to the factor Ambiguity, we limit ourselves to sentences in which the DP in the embedded clause is case-ambiguous. The ambiguity was derived by replacing a masculine DP in the embedded clause by a feminine, case-ambiguous counterpart. Our experimental design is laid out in Table 1:

Table 1: Experimental factors and levels + abbreviations

Factor		Ambiguity (Amb)	
		No Ambiguity (NoAmb)	**Embedded Ambiguity (EmbAmb)**
Argument (Arg)	**Subject (Sub)**	NoAmbSub	EmbAmbSub
	Object (Obj)	NoAmbObj	EmbAmbObj

3.1   Materials

Examples of the conditions are in (7) – (10) below.[3]

(7)   *Unambiguous subject question (NoAmbSub)*
       Welch-er   Schriftsteller denkst du,   dass den   Verleger   geschätzt   hat?
       Which-NOM   author   think you   that the.ACC publisher appreciated has
       'Which author do you think appreciated the publisher?'
(8)   *Unambiguous object question (NoAmbObj)*
       Welch-en Schriftsteller denkst du,   dass der   Verleger   geschätzt   hat?
       Which-ACC author   think you   that the.NOM publisher appreciated has
       'Which author do you think the publisher appreciated?'
(9)   *Embedded Ambiguity subject question (EmbAmbSub)*
       Welch-er   Schriftsteller denkst du, dass die   Verleger-in   geschätzt   hat?
       Which-NOM   author   think you that the.?   publisher-FEM   appreciated has
       'Which author do you think appreciated the (female) publisher?'
(10)   *Embedded Ambiguity object question (EmbAmbObj)*
       Welch-en Schriftsteller denkst du,   dass die   Verleger-in   geschätzt   hat?
       Which-ACC author   think you   that the.?   publisher-FEM   appreciated has
       'Which author do you think the (female) publisher appreciated?'

We created 24 items similar to the one above and divided these over three experimental lists in a Latin-square design, so that each participant saw only one particular lexicalization and eight sentences per condition. The full set of DPs and participles used to create the items are in the Appendix. In addition to the 48 experimental items, there were 16 filler items that involved LD subject and object questions with so-called 'embedded V2' clauses as in (11) and (12) below.[4]

(11)   Welch-er   Reporter denkst   du, hat   den   Filmstar gestört?
       Which-NOM   reporter think   you has   the.ACC   film.star disturbed
       'Which reporter do you think disturbed the film star?'

---

[3] In addition to these four conditions, there were two additional conditions in which the *wh*-phrase itself was case-ambiguous, which we will not report on here for reasons of space.

[4] Reis (1996) and Kiziak (2010) provide convincing arguments that these constructions do not involve proper embedding, but parenthesis, and therefore do not instantiate LD movement proper.

(12) Welch-en Häftling denkst du, hat der Wärter angeschrien?
Which-ACC prisoner think you has the.NOM guard yelled.at
'Which prisoner do you think the guard yelled at?'

The experimental items, as well as the filler items, were always followed by a comprehension task in which two statements appeared on the screen, one corresponding to a subject reading of the *wh*-phrase and one to an object reading of the *wh*-phrase, from which participants had to choose the correct reading. Example (13) shows the statements that appeared after the questions in (7) and (8), where (13a) would be the correct option for (7) and (13b) for (8):[5]

(13) *Comprehension statement*
a. Ich denke, dass der britische Schriftsteller den Verleger geschätzt hat
I think that the.NOM British author the.ACC publisher appreciated has
'I think the British author appreciated the publisher.'
b. Ich denke, dass der Verleger den britischen Schriftsteller geschätzt hat
I think that the.NOM publisher the.ACC British author appreciated has
'I think that the publisher appreciated the British author.'

3.2 Participants

Thirty native German speakers participated, all students at the University of Oldenburg. The data from one participant were excluded due to extremely slow reading times. The remaining 29 participants were aged 18–28 ($M = 23$, $SD = 3$). Twenty-two of them were female. Almost all participants came from Northern Germany, except for two Bavarian participants and two participants from Frankfurt.

3.3 Procedure

The experiment was conducted in a quiet room at the university using the software E-prime, version 2 (Psychology Software Tools 2012), and a Chronos response box. Participants first received oral and written instructions, explaining that they would see 64 sentences which they had to read as fast and accurately as possible. The task was a self-paced reading task in which the sentences appeared word-for-word (with the exception of DPs, which were presented as a whole) at the center of the screen. Each trial started with a fixation cross at the center of the screen, after which the participants could make the segments (words and DPs) appear by pressing a button on the response box. The duration between each button press was recorded. After each sentence was finished, two statements appeared on the center of the screen marked A and B, from which participants were asked to choose as fast as possible (see example 13). They were given a time limit of three seconds to answer; if they did not answer in time, a warning appeared on the screen and the next trial started. Before the experiment started, there was a practice round of ten sentences. All participants scored an accuracy of 70% or higher during this phase.

3.3 Analysis

Both comprehension data and reading time data were collected. The comprehension data were analyzed using generalized linear mixed-effects models in R. A full converging random effects structure was used, with random intercepts for participants and random slopes for Ambiguity and Argument over participants. Based on the research questions, Ambiguity and Argument were added as fixed effects, as well as an interaction between Ambiguity and Argument. Finally, the warranted inclusion of List and Trial number

---

[5] Because the experimental sentences used referentially specific 'which NP' *wh*-phrases, the statements (which were possible answers to the *wh*-questions) contained DPs modified by an adjective for reasons of felicity. The adjective always referred to a nationality.
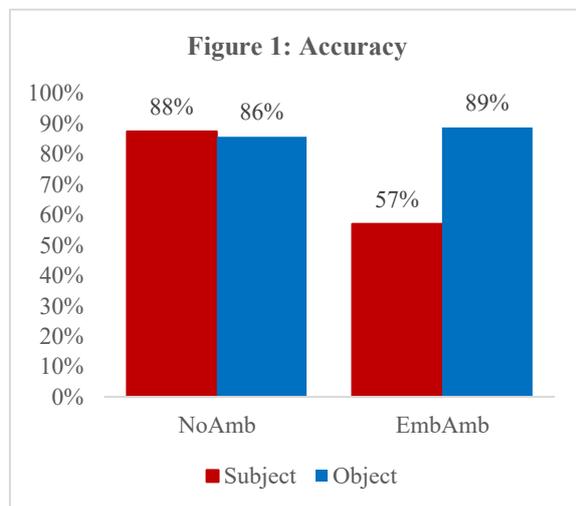
(sequential number in the experiment, centered) was tested based on model comparisons. Only Trial number improved the model and was therefore included as a fixed effect.

The reading time data from all (i.e. both correctly and incorrectly interpreted) trials from the self-paced reading task were first transformed to residual reading times, as these reflect normalized reading times based on word length and reading times of the rest of the sentence. These residual reading times for each segment (word or DP) were then analyzed separately using linear mixed-effects models in R. Full converging random effects structures were used, as well as the fixed factors of Ambiguity and Argument and an interaction between Ambiguity and Argument. Again, the warranted inclusion of List and Trial number was tested based on model comparisons, with Trial number improving all models and therefore being included as a fixed effect. When necessary, post-hoc Bonferroni-corrected pairwise comparisons were performed to better understand the effects obtained from the linear model. The effect of Trial number will not be discussed in the Results section, as it is not a factor of interest, but we can report that, unsurprisingly, participants read faster and answered more comprehension statements correctly over time.
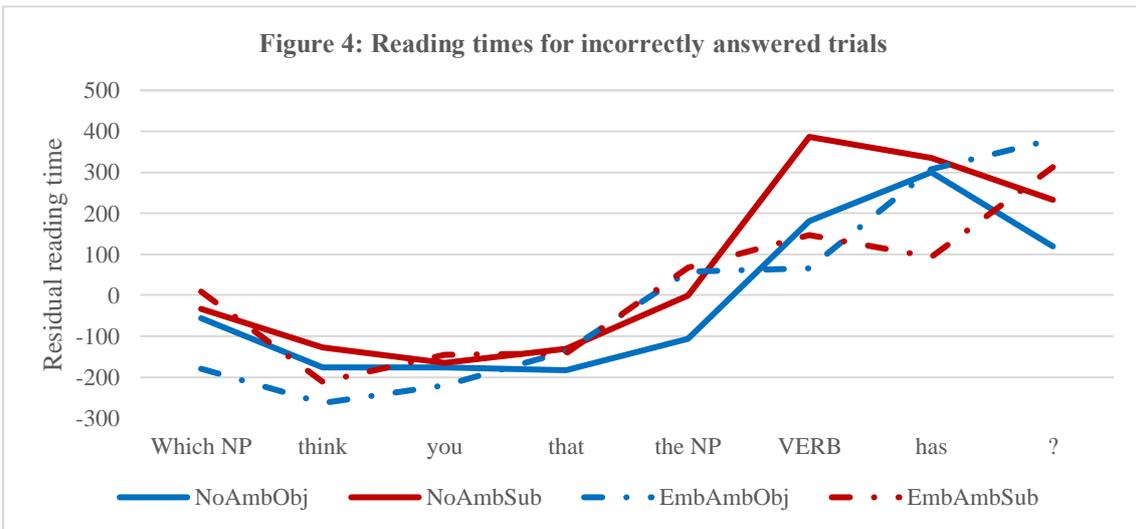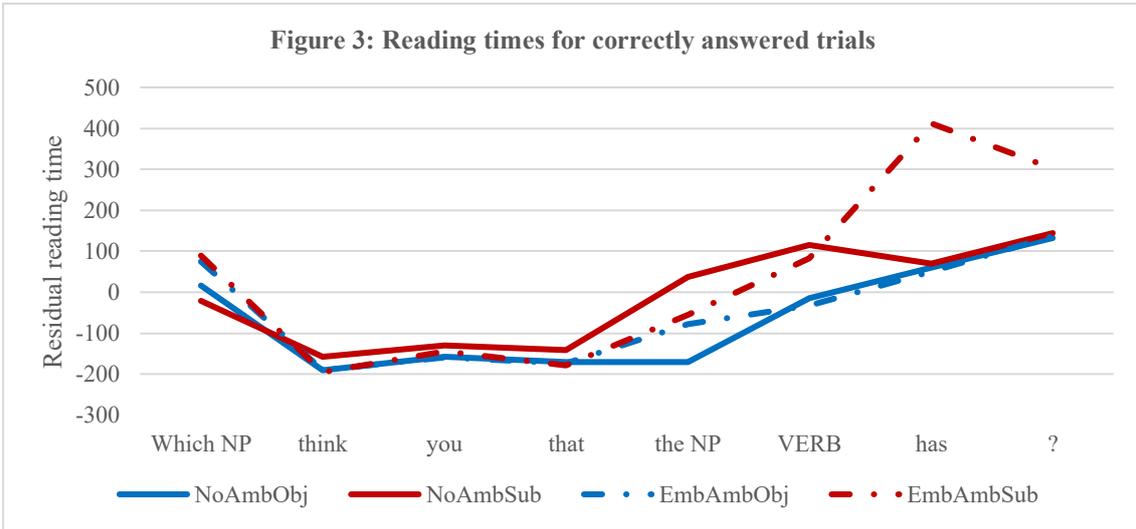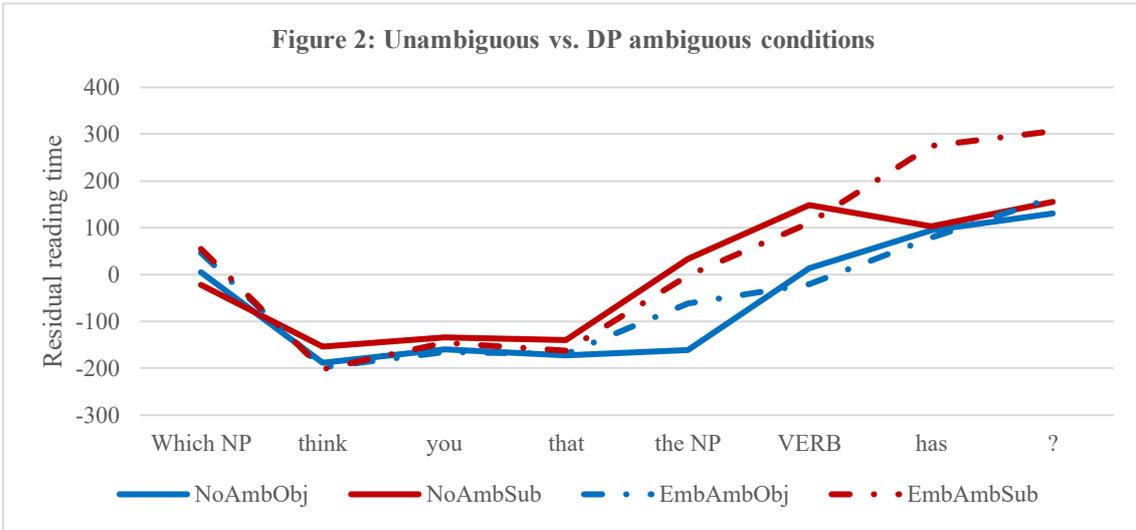
4. Results

The results for the comprehension questions are displayed in Figure 1. There were main effects of Argument ($p < 0.001$) and Ambiguity ($p < 0.001$), as well as a significant interaction between Argument and Ambiguity ($p < 0.001$): Participants were equally successful on NoAmb conditions, with accuracy $\geq 86\%$. For conditions with an embedded ambiguous DP, accuracy dropped to 57% for subject questions versus 89% for object questions.

The reading time data are depicted in Figure 2. As can be seen, reading times minimally differ throughout the main clause but start to diverge right after the complementizer. At the embedded DP, there is a significant effect of Argument ($p < 0.01$) and a significant interaction between Argument and Ambiguity ($p < 0.05$). Pairwise comparisons show

Figure 1: Accuracy

that this is due to the fact that the difference between subject and object questions is significant for NoAmb conditions ($p < 0.001$), but not for the EmbAmb conditions. At the following verb, there is a significant effect of Argument ($p < 0.01$), due to subject questions being read slower than object questions. At the auxiliary, there is a trend of an interaction between Argument and Ambiguity ($p = 0.05$), due to a significant difference between EmbAmb subject and object questions ($p < 0.05$). The differences in reading times at the sentence-final question mark did not reach significance.

Figure 2: Unambiguous vs. DP ambiguous conditions

Figure 3: Reading times for correctly answered trials

Figure 4: Reading times for incorrectly answered trials

It is customary in self-paced reading studies to include only reading time data of correctly answered trials; however, this did not fit the purpose of our study, as we were interested in the potential misinterpretation/incorrect analysis of LD subject questions. Nonetheless, it is informative to look at the reading times of correctly and incorrectly answered trials separately. These are depicted in Figures 3 and 4 respectively. Because of the large number of missing datapoints in both datasets, we did not perform any statistical analyses on these data. What Figure 3 and 4 show is that the overall pattern in Figure 2 corresponds more closely to the pattern of the correctly answered trials than to that of the incorrectly answered trials, in particular with respect to the EmbAmbSub condition.

5. Discussion

The central questions we sought to answer were whether German LD subject questions are misinterpreted as object questions and whether such misinterpretations occur more frequently in the context of local case ambiguities. Relatedly, we wanted to know if and when processing difficulties with LD subject questions occur. We discuss these questions in the following paragraphs.

5.1 'Good enough' processing in LD subject questions

Kiziak (2010) hypothesized that the COMP-trace effect is weaker in German because subject LD questions are sometimes misinterpreted as object questions, under the assumption that readers pursue 'good enough' processing strategies. We found evidence that readers indeed misinterpreted subject as object questions, but only in the context of a local case ambiguity. Thus, it is only in certain contexts that a weaker COMP-trace can be expected and explained as a result of 'good enough' processing, specifically in the presence of temporary case-ambiguities. The surprisingly low accuracy on EmbAmbSub questions shows that readers often go for an object reading of the *wh*-phrase, even though this clearly conflicts with the nominative case-marking on the *wh*-phrase. This strongly suggests that readers interpreted the embedded, case-ambiguous DP as the subject of the embedded verb, a reading that is locally possible but globally incorrect. The reading time data support this. For the unambiguous conditions, there is a significant difference between subject and object questions at the embedded DP, with subject questions being read more slowly. A possible explanation for this could be that the slowdown for subject questions is due to a type of superficial 'subject-first' preference in German, i.e. a preference for the first DP in the clause to be nominative. At least for temporary ambiguous constructions, such a preference has been shown to exist (see, amongst others, Meng & Bader 2000). However, such a subject-first preference for the embedded DP is somewhat unexpected in light of the fact that readers have previously encountered a nominative marked *wh*-phrase, which they were unable to integrate with the matrix verb. In this respect, it appears that the complementizer strongly reduces the expectation for an adjacent subject gap, causing a slowdown in reading times for the accusative DP directly following it. In other words, the slowdown at the embedded DP can be interpreted as an online reflection of the COMP-trace effect.

5.2 Misparses in LD subject questions

As the reading time data show, the case-ambiguous conditions do not show an asymmetry at the embedded DP: the residual reading times for subject and object questions are very close together and do not differ significantly. Furthermore, their reading times lie in between the reading times for the unambiguous conditions. This raises the question whether the EmbAmb questions pattern more with NoAmbSub or with NoAmbObj questions. In other words: are the EmbAmb conditions treated like a subject question or like an object question at the point of the embedded DP? We believe the latter option to be correct. For object questions, we see that the presence of an embedded ambiguous DP results in longer reading times (although the difference was not statistically significant in the post-hoc comparisons). There are two plausible explanations for this: first of all, in the EmbAmb conditions, the embedded DP was longer in terms of

characters, since it was derived from the DPs in the unambiguous conditions by adding the feminine suffix –*in*. Secondly, it is plausible that case-ambiguous DPs take longer to process than unambiguous ones regardless of their length. Thus, the slowdown for EmbAmbObj questions is likely not so much due to them being (temporarily) analyzed as subject questions, but simply due to a higher processing cost for ambiguous DPs. For EmbAmbSub questions, we would therefore expect the same relative slowdown at this position, but this was not the case, as was reflected by significant interactions between Argument and Ambiguity. This follows if we assume that EmbAmbSub questions are (at least temporarily) misparsed as object questions. This interpretation is also backed up by the comprehension data, which furthermore suggest that, in many cases, readers do not recover from this misparse. The reading time data at later positions further support this explanation: EmbAmbSub questions continue to be associated with higher reading times, reflected by a significant effect for Argument at the verb and a trend of an interaction between Argument and Ambiguity at the auxiliary, due to EmbAmbSub questions being read slower than all other conditions. Inspection of the reading times separated by correctly and incorrectly answered trials (Figures 3 and 4) suggest that, for correctly answered trials, there is a reanalysis effect for the EmbAmbSub condition: whereas reading times for EmbAmbSub and EmbAmbObj questions are close together at the embedded DP, they start to diverge at the verb and particularly at the auxiliary, which suggests that readers initially mistake an EmbAmbSub question for an object question but revise this later on. No such pattern is visible for the incorrectly answered trials; here the reading times between EmbAmbObj and EmbAmbSub questions only clearly diverge at the auxiliary, but in a different direction: EmbAmbSub questions are read faster at this position. In sum, both the comprehension data and the reading time data show that readers have problems with LD subject questions, reflected by longer reading times and comprehension problems in the presence of unambiguous case-marking in the embedded clause.

As we suggested at the end of the previous paragraph, encountering a complementizer appears to strongly reduce the expectation of an (adjacent) subject gap, even when readers have processed a nominative marked *wh*-phrase earlier that they haven't been able to integrate yet. Instead of postulating a subject gap immediately after the complementizer, readers frequently interpret a case-ambiguous embedded DP as the local subject. This suggests that they end up with a parse with two nominative DPs. Interestingly, previous research has shown that such double nominative constructions are frequently accepted. A speeded grammaticality judgement task by Schlesewsky and Frisch (2003) shows that participants are much less accurate in judging double nominative constructions compared to double accusative and double dative constructions, which they attribute to the unmarked, default status of nominative case. A speeded grammaticality judgment task by Schlesewsky et al. (2003) shows that the accuracy on judging the grammaticality of double nominative constructions decreases when the distance between the two DPs is increased, suggesting that comprehenders are more likely to ignore the case of the first DP when it's further away from the second DP. In our experiment, the DPs were in two different clauses, thus the linear as well as the structural distance was relatively large. Taken together, this could explain why readers relatively frequently tend to interpret the second DP as being nominative, even though this conflicts with the case of the initial nominative DP.

5.3   *The locus of processing problems in LD subject questions*

A final research question we formulated is whether the processing difficulty associated with LD subject questions is located in the main clause or in the embedded clause. It became clear that there are no significant differences in reading times in the main clause. This goes against Haider (2007) who has suggested that the lower acceptability of subject LD questions in German can be attributed to parsing problems in the main clause. Furthermore, the fact that effects do appear in the embedded clause, directly after processing the complementizer, strongly suggests that German LD subject questions do involve COMP-trace violations. This brings us back to the core questions concerning the COMP-trace effect, namely which constraint is responsible for the effect and why this constraint leads to ungrammaticality in one language (English) but degraded grammaticality in others (German). In the introduction, we discussed

the main syntactic analyses of the COMP-trace effect. We argued that it is difficult to account for the gradability of the effect under both freezing and anti-locality accounts, because under such accounts LD subject movement should either result in converging derivations and representations or it should not, but not something in between. In the next section, we will offer a possible solution for this conundrum.

5.4   COMP-trace effect and accessibility

We would like to propose the following: the COMP-trace effect should not be formulated in terms of an inviolable constraint. Rather, the problem with extracting subjects from embedded clauses should be phrased in terms of accessibility (cf. McDaniel et al. 2015). As McDaniel and colleagues point out, languages differ in the extent to which they allow extraction out of embedded clauses (i.e. non-local gaps). Based on Hawkins (2004), they propose the following hierarchy:

(14)   Clause embedding hierarchy for gaps:
            Infinitival complement > Finite complement > Subjacency structure

In addition to this hierarchy, there is a second hierarchy that relates to the syntactic function of non-local gaps. As McDaniel et al. point out, this hierarchy is the mirror image of the one that pertains to local gaps (i.e. within-clause movement; Keenan and Comrie's (1977) original Accessibility Hierarchy): for non-local gaps, subject gaps are less accessible than object gaps:

(15)   Hierarchy for non-local gaps: non-subjects > subjects

They illustrate this with various typological examples, and argue that the COMP-trace effect in English is just one reflection of this hierarchy.

The Accessibility Hierarchy for within-clause movement has been explained by assuming that lower accessibility is associated with more processing difficulty. Languages are known to differ in their cut-off points along the Accessibility Hierarchy, which can be explained by assuming languages differ as to whether they have grammaticalized processing constraints. This idea has been formalized in Hawkins's Performance-Grammar Correspondence Hypothesis (PGCH; Hawkins 2004). Analyzing the COMP-trace effect in these terms gives us a handle on why COMP-trace violations are dispreferred in both English and German, but only considered ungrammatical in English. For German, the PGCH explains why LD subject extraction is associated with more processing difficulty (as our data have shown) and lower acceptability (Featherston 2005; Kiziak 2010), without being ruled out altogether (Haider 2010). Furthermore, Hawkins argued that the PGCH is also reflected diachronically, in the sense that processing considerations can be the driving force behind grammatical change. In this respect, it is important to note that the COMP-trace effect has not always been active in English, but became effective around the first half of the 13th century (Bergh & Seppänen 1994). Thus, the relevant difference between English and German appears to be that English grammaticalized the lower accessibility of subjects for LD extraction at some point, whereas German did not.

Of course, viewing the COMP-trace effect in these terms raises an important question, which is why LD subject movement is more difficult to process than LD object movement. Hawkins (2004) gives an explanation for English which is based on the idea that complementizer deletion comes at a processing advantage for subject LD dependencies but not for object LD dependencies. Hawkins quantifies the complexity of an LD dependency by computing total 'domain differentials'. The domain differentials consist of the filler-gap domain (the *wh*-dependency), the lexical domain of the matrix and embedded verb (the respective verbs + their arguments), and the phrasal combination domain of the matrix verb (the matrix verb + the embedded clause). In case of subject LD dependencies, the embedded clause starts with a finite verb in the absence of a complementizer. According to Hawkins, finite verbs and complementizers can both do the same job, namely construct an embedded clause. Thus, the complementizer is completely superfluous in cases of LD subject movement, where the embedded clause starts with the finite verb in the

absence of a complementizer (or an adjunct phrase): it would only increase the length of the *wh*-dependency. In case of an object LD dependency, the embedded clause does not start with a finite verb (but with the subject). Therefore, from a processing perspective, it now becomes advantageous to pronounce the complementizer (which does make the *wh*-dependency one word longer, but makes early construction of the embedded clause possible, and therefore shortens the phrasal combination domain). Hawkins proposes that English has grammaticalized this processing advantage for complementizer deletion, making it obligatory. However, it is not directly clear how this explanation would carry over to German, since German differs in many relevant respects: it has a different word order (SOV instead of SVO), complementizer deletion results in verb fronting (to C), and it has overt case marking.

McDaniel et al. (2015), conversely, propose that the COMP-trace effect is not so much parsing-related, but motivated by production considerations. According to them, starting an embedded clause with a gap is problematic from a sentence-planning perspective. A way around this is to plan the matrix and the embedded clause simultaneously. McDaniel et al. argue that such simultaneous planning is reflected by complementizer deletion. With respect to the relationship between production and parsing, McDaniel et al. claim that "production and parsing principles might shape the grammar cooperatively, and in many cases it is difficult to tease the two apart" (McDaniel et al. 2015: 433). They refer to MacDonald and colleagues who have argued that production phenomena show up as parsing phenomena because parsing is affected by frequency of usage (Gennari & MacDonald 2009, MacDonald 2013, MacDonald & Thornton 2009, Race & MacDonald 2003). How does this translate to the COMP-trace effect that we established in our experiment? From a production perspective, the problem with a COMP-trace violation would be that an embedded clause starts with a gap. Such structures are therefore avoided in production and consequently infrequent. In terms of parsing, this would then have the effect that upon encountering a complementizer, the parser will avoid postulating a gap immediately adjacent to it. If an accusative DP directly follows the complementizer, this would signal that there is in fact a (subject) gap at the beginning of the clause. This explains the slow-down in terms of reading times for case-unambiguous subject questions. However, if a case-ambiguous DP follows the complementizer, this DP can function as a superficial work-around for having to process a subject gap, by interpreting this DP as the local subject.

In conclusion, our data as well as crosslinguistic patterns strongly suggest that the COMP-trace effect is best understood in terms of processing, which in many languages leads to these constructions being ruled out altogether. There is virtually no research on the processing of COMP-trace configurations. However, German, but also Dutch, are languages where this is possible, since COMP-trace violations are not categorically ruled out in these languages. The study reported here is just a first step towards this goal and we are aware of its limitations. Firstly, the number of items per participant (eight) was on the lower side. Secondly, we did not include a baseline condition (extraction without *dass*). Thirdly, in the ideal case, the comprehension and reading time data should be supplemented by acceptability judgment data. We are currently running follow-up experiments in our lab where these factors are taken into account and which we hope to report on in the future.

6. Conclusions

Although COMP-trace violations in German are not categorically ruled out, they are associated with decreased acceptability. The self-paced reading and comprehension data presented here showed that LD subject questions are also associated with higher processing difficulty and lead to frequent misinterpretations in the presence of local ambiguities. We argue that our data is best explained by assuming COMP-trace effects are in essence processing related, which has led to the exclusion of COMP-trace configurations in certain languages, including English.

References

Andersson, Sven-Gunnar and Sigmund Kvam, 1984. *Satzverschränkung im heutigen Deutsch*. Narr, Tübingen.

Meng, Michael and Markus Bader, 2000. Mode of disambiguation and garden-path strength: An investigation of subject-object ambiguities in German. *Language and Speech* 43 (1): 43–74. URL https://doi.org/10.1177/00238309000430010201.

Bayer, Josef. 2005. Was beschränkt die Extraktion?: Subjekt Objekt vs. Topic Fokus. In *Deutsche Syntax: Empirie und Theorie* edited by Franz-Joseph d'Avis, pp. 233–257. Acta Universitatis Gothoburgensis, Göteborg.

Bayer, Josef and Martin Salzmann, 2013. That-trace effects and resumption: How improper movement can be repaired. In *Repairs: The Added Value of Being Wrong*, edited by Patrick Brandt and Eric Fuss, pp. 275 - 333. Walter de Gruyter, Berlin. URL https://doi.org/10.1515/9781614510796.

Bennis, Hans, 1986. *Gaps and Dummies*. Foris, Dordrecht. URL https://doi.org/10.1515/9783110889536.

Bergh, Gunnar and Aimo Seppänen, 1994. Subject Extraction in English: The Use of the That-complementizer. In *English Historical Linguistics 1992: Papers from the 7th International Conference on English Historical Linguistics, Valencia, 22–26 September 1992* edited by Francisco Fernández, Miguel Fuster Márquez and Juan Jose Calvo, pp. 131–143. John Benjamins Publishing Company, Amsterdam/Philadelphia. URL https://doi.org/10.1075/cilt.113.15ber.

Boeckx, Cedric, 2008. *Bare syntax*. Oxford University Press, Oxford.

Bošković, Željko, 2016. On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. *The Linguistic Review* 33 (1): 17–66. URL https://doi.org/10.1515/tlr-2015-0013.

Bresnan, Joan. 1972. *Theory of complementation in English syntax*. Ph.D: thesis, Massachusetts Institute of Technology, Cambridge, MA. URL https://doi.org/10.4324/9781315458731.

Brillman, Ruth and Aron Hirsch, 2016. An anti-locality account of English subject/non-subject asymmetries. In *Proceedings from the 50th Annual Meeting of the Chicago Linguistic Society*, edited by Ross Burkholder, Carlos Cisneros-Juarez, Emily Coppess, Julian Grove, Emily Hanink, Hilary Head McMahan, Cherry Meyer, Natalia Pavlou, Özge Sarigül, Adam Roth Singerman and Anqi Zhang. Chicago Linguistic Society, Chicago.

Chomsky, Noam, 1981. *Lectures on government and binding*. Foris, Dordrecht. URL https://doi.org/10.1515/9783110884166.

Chomsky, Noam, 1986. *Barriers*. MIT Press, Cambridge, MA.

Chomsky, Noam and Howard Lasnik, 1977. Filters and control. *Linguistic Inquiry* 8 (3): 425–504.

Cole, Peter and Gabriella Hermon, 1981. Subjecthood and Islandhood: Evidence from Quechua. *Linguistic Inquiry* 12: 1–30.

Culicover, Peter, 1993. Evidence against ECP accounts of the that-t effect. *Linguistic Inquiry* 24: 557–561. URL https://doi.org/10.1093/acprof:oso/9780199660230.003.0009.

Douglas, Jamie, 2017. Unifying the that-trace and anti-that-trace effects. *Glossa: A Journal of General Linguistics* 2 (1): 60. URL http://doi.org/10.5334/gjgl.312.

Fanselow, Gisbert and Stefan Frisch, 2006. Effects of processing difficulty on judgments of acceptability. In *Gradience in grammar* edited by Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky and Ralf Vogel, pp. 291–316. Oxford University Press, Oxford. URL https://doi.org/10.1093/acprof:oso/9780199274796.003.0015.

Fanselow, Gisbert, Reinhold Kliegl and Matthias Schlesewsky, 1999. Processing difficulty and principles of grammar. In *Constraints on language: Aging, grammar, and memory* edited by Gisbert Fanselow, Reinhold Kliegl, Matthias Schlesewsky and Susan Kemper, pp. 170–200. Springer, Boston, MA. URL https://doi.org/10.1007/0-306-46902-2_7.

Featherston, Sam, 2005. That-trace in German. *Lingua* 115: 1277–1302. URL https://doi.org/10.1016/j.lingua.2004.04.001.

Ferreira, Fernanda and Nicole D. Patson, 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1 (1–2): 71–83. URL https://doi.org/10.1111/j.1749-818X.2007.00007.x.

Gennari, Silvia and Maryellen MacDonald, 2009. Linking production and comprehension processes: The case of relative clauses. *Cognition* 111: 1–23. URL https://doi.org/10.1016/j.cognition.2008.12.006.

Ha, Seungwan, 2010. A prosody analysis of the comp-trace effect. *English Language and Linguistics* 16 (2): 109–133. URL https://doi.org/10.17960/ell.2010.16.2.005.

Haegeman, Liliane, 1983. *Die* en *dat* in West-Flemish relative clauses. In *Linguistics in the Netherlands* edited by Hans Bennis & Wus van Lessen Kloeke, pp. 83–91. Assen, Van Gorcum.

Haider, Hubert, 2007. As a matter of facts–comments on Featherston's sticks and carrots. *Theoretical Linguistics* 33 (3): 381–394. URL https://doi.org/10.1515/TL.2007.025.

Haider, Hubert, 2010. *The syntax of German*. Cambridge University Press, Cambridge. URL https://doi.org/10.1017/CBO9780511845314.

Hawkins, Jack A., 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75: 244–285. URL https://doi.org/10.2307/417261.

Hawkins, Jack A., 2004. *Efficiency and Complexity in grammars*. Oxford University Press, Oxford. URL https://doi.org/10.1093/acprof:oso/9780199252695.001.0001.

Hermon, Gabriella, 1984. *Syntactic modularity*. Foris, Dordrecht. URL https://doi.org/10.1515/9783110849141.

Ishii, Toru, 2004. The Phase Impenetrability Condition, the Vacuous Movement Hypothesis, and That-t Effect. *Lingua* 114: 183–215. URL https://doi.org/10.1016/S0024-3841(03)00045-7.

Jackson, Carrie N., and Susan C. Bobb, 2009. The processing and comprehension of *wh*-questions among second language speakers of German. *Applied Psycholinguistics* 30 (4): 603–636. URL https://doi.org/10.1017/S014271640999004X.

Kandybowicz, Jason, 2006. Comp-Trace effects explained away. In *Proceedings of the 25th West Coast Conference on Formal Linguistics* edited by Donald Baumer, David Montero and Michael Scanlon, pp. 220–228. Cascadilla Proceedings Project, Somerville, MA.

Kandybowicz, Jason, 2008. *The Grammar of Repetition: Nupe grammar at the syntax-phonology interface*. John Benjamins Publishing Company, Amsterdam/Philadelphia. URL https://doi.org/10.1075/la.136.

Keenan, Edward L. and Bernard Comrie, 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8 (1): 63–99.

Kiziak, Tanja, 2010. *Extraction Asymmetries: Experimental evidence from German*. John Benjamins Publishing Company, Amsterdam/Philadelphia. URL https://doi.org/10.1075/la.163.

Lasnik, Howard and Mamoru Saito, 1984. On the nature of proper government. *Linguistic Inquiry* 15: 235–289. URL https://doi.org/10.1007/978-94-015-6859-3_10.

Lasnik, Howard and Mamoru Saito, 1992. *Move α*. MIT Press, Cambridge, MA.

MacDonald, Maryellen C., 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4: 1–16. URL https://doi.org/10.3389/fpsyg.2013.00226.

MacDonald, Maryellen C. and Robert Thornton, 2009. When language comprehension reflects production constraints: Resolving ambiguities with the help of past experience. *Memory and Cognition* 37: 1177–86. URL https://doi.org/10.3758/MC.37.8.1177.

McDaniel, Dana, Cecile McKee, Wayne Cowart and Merrill F. Garrett, 2015. The role of the language production system in shaping grammars. *Language* 91: 415–441. URL https://doi.org/10.1353/lan.2015.0021.

McFadden, Thomas and Sandhya Sundaresan, 2017. Unifying EPP and COMP-trace effects: constraints on silent elements at the edge. *Glossa: A Journal of General Linguistics* 3 (1): 43. URL https://doi.org/10.5334/gjgl.419.

Müller, Gereon and Joachim Sabel, 1989. Aufgebrochene Barrieren. *Frankfurter Linguistische Forschungen* 6: 22–40.

Perlmutter, David, 1971. *Deep and Surface Structure Constraints in Syntax*. Holt, Rinehart, and Winston, New York.

Pesetsky, David, 1982. Complementizer-Trace Phenomena and the Nominative Island Condition. *The Linguistic Review* 1: 297–345. URL https://doi.org/10.1515/tlir.1982.1.3.297.

Pesetsky, David and Esther Torrego, 2001. T-to-C Movement: Causes and Consequences. In *Ken Hale: A Life in Language* edited by Michael Kenstowicz, pp. 355–426. MIT Press, Cambridge, MA. URL https://doi.org/10.7551/mitpress/4056.003.0014.

Psychology Software Tools, 2012. *E-Prime 2.0*. Pittsburgh, PA.

Race, David S. and Maryellen C. MacDonald, 2003. The use of 'that' in the production and comprehension of object relative clauses. In *Proceedings of the 25th annual conference of the Cognitive Science Society* edited by Richard Alterman and David Hirsh, pp. 946–951. Psychology Press, New York. URL https://doi.org/10.4324/9781315799360.

Reis, Marga, 1996. Extractions from verb-second clauses in German? In *On Extraction and Extraposition in German* edited by Uli Lutz and Jürgen Pafel, pp. 45–88. John Benjamins Publishing Company, Amsterdam/Philadelphia. URL https://doi.org/10.1075/la.11.03rei.

Rizzi, Luigi and Ur Shlonsky, 2007. Strategies of subject extraction. In *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics* edited by Uli Sauerland and Hans-Martin Gartner, pp. 115–160. Walter de Gruyter, Berlin. URL https://doi.org/10.1515/9783110207552.115.

Rizzi, Luigi, 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Sato, Syuhei and Yoshinori Dobashi, 2016. Prosodic phrasing and the that-trace effect. *Linguistic Inquiry* 47: 333–349. URL https://doi.org/10.1162/LING_a_00213.

Schippers, Ankelien, 2012. Variation and change in Germanic long-distance dependencies. Ph.D. thesis, University of Groningen, Groningen. Retrieved from http://www.rug.nl/research/portal/files/2459573/dissertation_final_version.pdf.

Schlesewsky, Matthias and Stefan Frisch, 2003. Nominative case as a multidimensional default. In *Linguistics in Potsdam 21* edited by Susann Fischer, Ruben van de Vijver and Ralf Vogel, pp. 61–72.

Schlesewsky, Matthias, Gisbert Fanselow, Reinhold Kliegl and Josef Krems, 2000. The subject preference in the processing of locally ambiguous *wh*-questions in German. In *German sentence processing* edited by Barbara Hemforth and Lars Konieczny, pp. 65–93. Springer, Dordrecht. URL https://doi.org/10.1007/978-94-015-9618-3_3.

Schlesewsky, Matthias, Gisbert Fanselow and Stefan Frisch, 2003. Case as a trigger for reanalysis: Some arguments from the processing of double case ungrammaticalities in German. In *Linguistics in Potsdam 21* edited by Susann Fischer, Ruben van de Vijver and Ralf Vogel, pp. 31–60.

Sobin, Nicholas, 1987. The variable status of Comp-trace phenomena. *Natural Language & Linguistic Theory* 5 (1): 33–60. URL https://doi.org/10.1007/BF00161867.

Appendix

Item	DP1	DP2	Participle
1	Schriftsteller *author*	Verleger *publisher*	geschätzt *appreciated*
2	Mörder *killer*	Polizist *officer*	bedroht *threatened*
3	Manager *manager*	Pianist *pianist*	respektiert *respected*
4	Minister *minister*	Sekretär *secretary*	gewarnt *warned*
5	Chirurg *surgeon*	Pfleger *nurse*	gelobt *praised*
6	Sänger *singer*	Chorleiter *choir director*	bewundert *admired*
7	Winzer *winemaker*	Gastronom *caterer*	kontaktiert *contacted*
8	Experte *expert*	Direktor *director*	getroffen *met*
9	Kellner *waiter*	Gastgeber *host*	geärgert *annoyed*
10	Student *student*	Professor *professor*	verteidigt *defended*
11	Konkurrent *competitor*	Sportler *sportsman*	getreten *kicked*
12	Maler *painter*	Galerist *gallerist*	begrüßt *greeted*
13	Musiker *musician*	Dirigent *conductor*	gerufen *called*
14	Investor *investor*	Mieter *tenant*	verärgert *annoyed*
15	Richter *judge*	Anwalt *lawyer*	befragt *questioned*
16	Physiker *physicist*	Philosophen *philosopher*	attackiert *attacked*
17	Politiker *politician*	Bürger *citizen*	provoziert *provoked*
18	Tourist *tourist*	Anwohner *local*	beleidigt *offended*