

The Effect of Clause Type on Long Extraction in German

Vasiliki Koukouloti & Markus Bader

University of Frankfurt

vasiliki.koukouloti@gmail.com, bader@em.uni-frankfurt.de

1 Introduction

Long extraction (LE for short) refers to the movement of a phrase over a clausal boundary. In English and other languages, object extraction out of a *that*-clause is accepted as grammatical, whereas subject extraction is not acceptable – a phenomenon known as the *that*-trace effect. LE in general, and the existence of a subject-object asymmetry, is a controversial issue in the syntax of German. While some researchers claim that the presence of a subject-object asymmetry is subject to dialectal variation (Fanselow 1987; Grewendorf 1995), others contend that a subject-object asymmetry is not present in German (Lühr 1988) and still others that it is rather a transfer of English judgments to German (Müller & Sabel 1989).

Starting with Andersson & Kvam (1984), LE in German has been a continuous topic of experimental research (e.g., Featherston 2005; Kiziak 2010). The central question addressed by most of this work concerns the existence and the status of the *that*-trace effect. Both Featherston (2005) and Kiziak (2010) found an asymmetry between subjects and objects concerning extractability from *that*-clauses. The overall acceptability of LE as well as the magnitude of the subject-object asymmetry varied as a consequence of certain experimental manipulations. For example, LE into SpecCP of a main clause question (long *wh*-movement) was more acceptable than LE into SpecCP of a declarative main clause (long topicalization), and the subject-object asymmetry was larger in the former case than in the latter.

LE is of particular interest because diachronic evidence suggests that it was quite common in earlier stages of German, but was then replaced to a large extent by alternative syntactic means (Paul 1916/19; for Dutch, cf. Hoeksema & Schippers 2012). Nevertheless, authentic examples are still easily found by searching the web, as shown in (1).

- (1) *Wen dachtet ihr, dass wir da zugelost bekommen? Arminia Bielefeld?*
who thought you that we there assigned get Arminia Bielefeld
'Who did you think that we get assigned there? Arminia Bielefeld?'

(<http://s.mitpicke.de/ucldraw-c0b4d79d3643>, last access July 13, 2017)

In addition to the subject-object asymmetry, two further issues arise. The first one concerns the potential effect of the landing site. The asymmetry between long *wh*-movement and long topicalization raises the question whether such an asymmetry exists also among embedded clauses. Despite the general interest in LE, LE into

embedded clauses has been hardly addressed in experimental studies. To fill this gap, Bader & Koukouloti (in prep.) have run a series of experiments that have investigated whether the landing site of the extracted element affects the acceptability of LE, focusing on the acceptability of LE into embedded clauses. The second issue addressed in our research concerns the potential effect of the availability of alternative structures on the acceptability of LE.

In this paper, we go a step further and present a new experiment that compares the acceptability of LE in all four constructions that – following Chomsky (1977) – are usually claimed to involve A'-movement. In particular, we compare two structures in which the landing site is within a main clause (comparative clauses and long topicalizations) and two structures in which the landing site is within an embedded clause (embedded questions and relative clauses). Interestingly, embedded questions and relative clauses differ with respect to the existence of an alternative structure: for the embedded wh-questions there is not any alternative, whereas for the relative clauses there is one (for details see below). In this way we can shed light on the issue of the existence of an alternative structure.

Before presenting the experiment, we briefly review previous studies on LE into main clauses in Section 2. We then report some observations from an ongoing corpus study in Section 3, showing that LE into embedded clauses can be found in Modern German. In Section 4, we summarize some of the results obtained by Bader & Koukouloti (in prep.) with regard to extraction into an embedded question and into a relative clause. This sets the background for the new experiment, which is presented in Section 5. The paper concludes with a general discussion in Section 6.

2 Previous Studies

The majority (if not all) of the studies on LE are related to the issue of the subject-object extraction asymmetry. On the basis of corpus data and grammaticality judgments, Andersson & Kvam (1984) suggested that a subject-object asymmetry is detectable in Modern German, although with some restrictions, which relativize the effect. The grammaticality judgments of Andersson & Kvam were elicited on the basis of a small number of specific examples and not in strict experimental settings, making generalizations difficult. Nevertheless, the characterization of LE in German for which they argued, has stood the test of time. For example, Featherston (2005) provides evidence for a that-trace effect for long wh-movement and long topicalization using magnitude estimation. Regarding LE into wh-questions, he also shows that the pattern is similar to previous findings for English, a language with undisputable that-trace effect.

Kiziak (2010) replicates both the effect of clausal type as well as the subject-object asymmetry using the thermometer judgment task, which is a variant of magnitude estimation. In particular, the judgments for topicalization were overall worse than those for wh-movement. Moreover, the difference between subject and object extraction was smaller for topicalization than for wh-movement, as in the former there were floor effects. The effect of clausal type could be due to the absence of a context, which would make the topicalization natural, a major drawback of Kiziak's study, as she herself points out. Thus, the issue of context has to be addressed before drawing conclusions on LE.

Another line of research has investigated the effect of parsing strategies on the subject-object extraction asymmetry. Fanselow & Frisch (2006) have shown that the acceptability of subject LE into wh-questions is affected by the local well-formedness of the sentences. They tested subject LE into wh-questions like the following:

- (2) *Was glaubst du (...)?*
 what believe you
 'What do you believe (...)?'
- (3) **Wer glaubst du (...)?*
 who believe you
 'Who do you believe (...)?'

The sequence in (2) is locally perfectly well-formed, as the w-pronoun *was* is case-ambiguous and it can be parsed as the object of the matrix verb. The sequence in (3), in contrast, is ungrammatical, as *wer* is case-unambiguous and it cannot be parsed as subject or as object of the matrix verb. Sentences with the ambiguous pronoun were significantly more acceptable than those with unambiguous pronouns. The interpretation is that local well-formedness affects global acceptability. Similar results are reported by Kiziak (2010), who tested wh-DPs instead of wh-pronouns and extraction from sentences with a case-ambiguous internal DP. We will come back to this effect of well-formedness later in our discussion of the results.

All in all, previous studies have shown that the controversial subject-object asymmetry can be detected in German, although modulated by structural (clausal asymmetry) and parsing factors (local well-formedness).

3 Some Observations Concerning LE in Written Language

Before starting with our experimental investigations, we tried to determine whether instances of LE can be found in authentic texts. Because we did not find relevant examples in even large corpora, like the deWac corpus (Baroni et al. 2009), we used the search facility provided by Google instead. As already shown by example (1), instances of LE can be found in the web. At the moment, we do not yet have quantitative evidence and will therefore confine ourselves to some observations we have made when searching for LE.

First of all, there are also instances of LE into embedded clauses. (4) is an instance of LE into an embedded question.

- (4) *Bevor ich mich verabschiede, frage ich Booger noch,*
 before I myself say goodbye ask I Booger yet
was er meint, dass am Dorf verbesserungswürdig wäre.
 what he thinks that at village in need of improvement would be

'Before I say goodbye, I ask Booger what he thinks has to be improved in the village.'

(<http://lotgdforum.ws/1223-die-gr-ne-ausgabe-93.html>, last access July 13, 2017)

(5) and (6) show instances of LE where a w-relative pronoun has been moved.

- (5) *Mein Dad soll alles anzeichnen, was er meint, daß raus muß.*
my dad shall everything mark what he thinks that out must
'My dad should mark everything that he thinks has to be removed.'

<http://www.landtreff.de/sehr-spate-erstdurchforstung-aber-wie-jetzt-richtig-t70643-30.html>, last access July 13, 2017)

- (6) *Dort bietet er weiterhin das an, was er hofft, dass Schramberg braucht.*
there offers he still that PRT what he hopes that Schramberg needs
'There, he is still offering what he thinks Schramberg needs.'

<http://www.schwarzwaelder-bote.de/inhalt.schramberg-firma-sporthug-treffen-die-nachwehen-der-insolvenz.aaf99816-7b81-43ee-888e-478d62fe8d97.html>,
last access July 13, 2017)

We did not find instances of LE proper with d-relative pronouns, although we found a couple of examples with a resumptive pronoun, as in (7).

- (7) *Sei ehrlich, wenn du etwas gefragt wirst, denn wenn du nur
be truthful when you something ask are then when you only
die Antworten gibst, die du glaubst, dass der Vermieter sie hören möchte:
the answers give which you think that the landlord them hear likes*

Er wird es merken!
he will it notice

'Say the truth when you are asked something, because, if you only give those answers of which you think that the landlord will hear them: He will notice it.'

<http://www.gutefrage.net/frage/hat-euch-die-wohnungssuche-auch-gefrustet-hattet-ihr-schwierigkeiten-eine-wohnung-zu-finden-wie-hat-man-am-meisten-chancen-eine-wohnung-zu-bekommen>, last access July 13, 2017)

Overall, it is our impression that authentic examples of LE in written language are either from temporary colloquial German, as found in informal internet bulletin boards, or from older stages of German (in particular from the 19th century). Paul cites many examples of LE structures. (8), (9) and (10) are examples of LE into main questions, embedded questions and relative clauses, respectively (all taken from Paul 1916/19).

- (8) *Was meynst du wohl, daß ich da sah*
what think you PRT that I there saw
'Well what do you think that I saw there.'

(Lenz, Lustp., 140)

- (9) *Sage mir doch, was du glaubtest, das ich von einem Bäre an mir hätte.*
tell me PRT what you believed that I of a bear at me had.
'But tell me, what do you think, that I had of a bear '

(E. Schlegel 52, 10)

- (10) *die Erklärung, die du willst, das ich geben soll.*
the explanation that you want that I give shall.
'the explanation that you want me to give to you'

(Schiller Br. 2, 203)

Examples from websites adhering to a formal style (e.g., newspapers) can also be found sometimes, as seen in (6), but their frequency seems to be low. This concurs with claims in the literature that LE was replaced in Modern German by other constructions and, thus, marginalized. For example, an alternative structure to long relativization is the *von* + relative pronoun construction in (11).

- (11) *Sie ist ein Kind der Zeit, nimmt sich alles, von dem sie meint, dass es ihr zusteht.*
 She is a child the time, takes herself everything, of which she thinks, that it her appertains.

‘She is a child of time, she takes everything what she thinks that is due to her’

(<https://leserunden.de/index.php?topic=11530.10;wap2>, last access July 13, 2017)

However, in less formal styles – especially in spoken language, but also in written registers closer to spoken language – LE still seems to be in use. An exception appears to be the verb *wollen* (‘want’). LE with this verb still enjoys some popularity even in formal registers in Modern German, especially in business contexts as in the following example.

- (12) *Was wollen Sie, dass Ihre Marke darstellt?*
 what want you that your brand represents
 ‘What do you want that your brand represents?’

(<http://www.swmt.org/dokumentationen/2012.html>, last access July 13, 2017)

4 LE into an Embedded Question

As mentioned above, the vast majority of studies on LE out of *dass*-clauses has focused on the asymmetry between subject and object movement into main clauses. Not much is known about LE when the landing site is contained within an embedded clause. To remedy this situation, Bader & Koukouloti (in prep.) ran a series of experiments investigating embedded questions and relative clauses. We summarize some of these experiments in this section in order to provide the background for our new experiment presented in Section 5.

4.1 LE into an Embedded Question

As shown by the corpus examples discussed in the preceding section, LE targets both main and embedded clauses. With regard to the availability of alternative syntactic means, there is an important difference between direct and indirect questions. LE of a *wh*-phrase into a main clause question has a fully acceptable alternative in Standard German, namely extraction from an embedded V2-clause (see (13a) and (13b), respectively).

- (13) a. *Was meinst du, dass Maria einkaufen soll?*
 what think you that Maria buy should
 ‘What do you think that Maria should buy?’

- b. *Was meinst du, soll Maria einkaufen?*
 what think you should Maria buy
 'What do you think Maria should buy?'

An alternative of this kind is not available for embedded questions, as witnessed by the ungrammaticality of (14b).

- (14) a. *Hast du Maria gefragt, was sie meint, dass wir benutzen können?*
 have you Maria asked what she thinks that we use can
 'Did you ask Maria what she thinks that we can use?'
- b. **Hast du Maria gefragt, was sie meint, können wir benutzen?*
 have you Maria asked what she thinks can we use
 'Did you ask Maria what she thinks we can use?'

Since the decline of LE in grammar and use has been attributed to the rise of alternative constructions (Paul 1916/19), the availability of an alternative structure for main clause questions but not for embedded questions could result in a difference in acceptability. If so, the extraction of a wh-phrase into a main clause question should be less acceptable than the extraction into an embedded question.

In contrast to Standard German, colloquial German provides an alternative to LE for both main and embedded clauses, namely partial movement. Sentences with partial movement are those in which the w-pronoun has not moved the whole way up to SpecCP of the question clause (see (15a) and (15b) for examples of partial movement in main and embedded questions, respectively).

- (15) a. [_{CP2} *Was meinst du*, [_{CP1} *was Maria einkaufen soll*]]?
 what think you what Maria buy should
 'What do you think that Maria should buy?'
- b. *Hast du Maria gefragt, [_{CP2} *was sie meint*, [_{CP1} *was wir benutzen können*]]?*
 have you Maria asked what she thinks that we use can
 'Did you ask Maria what she thinks (that) we can use?'

To test whether LE is judged differentially in main clause questions than in embedded questions, we compared sentences as in (13a) to sentences as in (14a). Partial movement sentences served as control sentences.

- (16) *Du kennst Klaus doch gut.*
 'You know Klaus well.'
- a. LE in embedded question
Du könntest fragen, was er meint, dass uns aus dieser Situation retten könnte.
 you could ask what he thinks that us out this situation rescue could
 'Perhaps you could ask him what he thinks could rescue us out of this situation?'
- b. LE in main clause question
Was meint er, dass uns aus dieser Situation retten könnte?
 what thinks he that us out this situation rescue could
 'What does he think could rescue us out of this situation?'

Our participants, who were all students from the University of Frankfurt and native speakers of German, had to judge the acceptability of such sentences on a Likert scale ranging from 1 (totally unacceptable) to 7 (totally acceptable). The results did not show any effect of embedding.¹ Main clause questions and embedded questions did not differ with regard to the acceptability of LE. LE was of marginal acceptability in both cases (about 3 on the 1-to-7 Likert scale). Control sentences with partial movement were scored as highly acceptable, again with no difference between main and embedded questions (main acceptability score around 6). Thus, the existence in Standard German of an alternative structure for LE into main clause questions does not reduce their acceptability. Moreover, these data provide preliminary evidence that extraction into embedded clauses is not different from extraction into main clauses and, consequently, that it does not matter for acceptability whether the landing site is contained within a main clause or within an embedded clause.

4.2 LE into a Relative Clause

LE into a relative clause is usually considered as ungrammatical in Modern German (e.g., Andersson & Kvam 1984; Lühr 1988), although it seems to have been acceptable in earlier times, as witnessed by the examples collected in Paul (1916/19). However, as shown above in examples (5) and (6), authentic examples can still be found. To assess the status of LE into a relative clause, we ran an experiment comparing relative clauses – all introduced by the *wh*-pronoun *was* ('what') – to corresponding indirect questions. Example sentences illustrating the design of this experiment are shown in (17) (embedded questions) and (18) (relative clauses).

(17) *Max sagte mehrfach, was er glaubt, dass den Kindern gefallen könnte.*
Max said several times what he thinks that the children please could
'Max said several times what he thinks could please the children.'

(18) *Max kaufte alles, was er glaubt, dass den Kindern gefallen könnte.*
Max bought all what he thinks that the children please could
'Max bought everything that he thought could please the children.'

We used two different judgment tasks in this experiment – judgments on a Likert scale from 1 to 7 and binary grammaticality judgments. The major finding was that with both tasks LE into a relative clause is somewhat more acceptable than LE into an embedded question. LE into embedded questions received a mean acceptability of 4.0 versus 4.7 for LE into relative clauses in the Likert scale acceptability task. LE into embedded questions were judged as grammatical in 50% of the cases versus 59% for LE into a relative clause in the binary grammaticality judgment task. Control sentences received substantially higher ratings, but with the same difference between embedded questions and relative clauses, which indicates that embedded questions were somewhat less acceptable in this experiment than relative clauses, independently of the presence of LE. What we can conclude in any case is that LE into a relative clause is at least as acceptable as LE into an embedded question. Thus, contrary to claims in the literature, LE into relative clauses is not banned from Modern German.

¹ All differences that we report in this section are statistically significant.

If the acceptability of LE depends only on the clause containing the landing site, then there should not be any difference between relative clauses with a w-pronoun, as in the previous experiment, and those with a d-pronoun. However, in our corpus search, we only found examples with w-pronouns. Given the difficulties of finding relevant examples, we cannot exclude that examples with d-pronouns also occur, but that we did not find them. We therefore compared the two types of relative clauses in an acceptability judgment task using again a 1-to-7 Likert scale. The material consisted of sentences as in (19) and (20).

- (19) *Rainer hat jetzt **ein Buch** bestellt, **das** er glaubt,*
 Rainer has now a book ordered that he thinks

dass man seinem Vater schenken könnte.
 that one to his father present could

‘Rainer has ordered a book that he thinks, that someone could present to his father.’

- (20) *Rainer hat jetzt **etwas** bestellt, **was** er glaubt, dass man seinem*
 Rainer has now something ordered that he thinks that one to his

Vater schenken könnte.
 father present could

‘Rainer has ordered something that he thinks, that someone could present to his father.’

The results confirmed to some extent the corpus data. W-relative clauses had a higher average acceptability rate than d-relative clauses (4.3 vs. 3.5). The difference between the two types of relative clauses is not large, however. Moreover, d-relative clauses with an average acceptability rate of 3.5 do not seem to be categorically excluded in German. There is one more point that makes these findings inconclusive concerning the difference between w- and d-relative clauses. The antecedent of the relative pronoun in each experimental condition had a different word class status. In the d-relative clause, the antecedent was a definite or indefinite NP; in the w-relative clause, it was a quantificational pronoun (*etwas* or *alles*). Although the antecedent NPs were selected in such a way that they constitute typical antecedents of each relative clause pronoun, it might be the case that this difference in the word class status blurs the results. Moreover, the definiteness of the noun antecedent, which was a between items factor, could also be a confounding factor.

In a follow-up experiment, we compared again d- and w-relative clauses, this time with the indefinite pronoun *etwas* as antecedent in both cases. According to Brandt & Fuß (2014), the indefinite pronoun *etwas* is equally compatible with *das* and *was*. The sequences *etwas, was* and *etwas, das* are comparably frequent and the Duden grammar regards the relativizers *das* and *was* as interchangeable in combination with *etwas*. The material was similar to the previous experiment, see examples (21) and (22).

- (21) *Rainer hat jetzt etwas bestellt, das er glaubt, dass man*
 Rainer has now something ordered that he thinks that one

seinem Vater schenken könne.
 to his father present could

‘Rainer has now ordered something that one could present to his father.’

- (22) *Rainer hat jetzt etwas bestellt, was er glaubt, dass man
Rainer has now something ordered that he thinks that one
seinem Vater schenken könne.
to his father present could.*

‘Rainer has now ordered something that one could present to his father.’

The acceptability pattern of the previous experiment appears in this one too. Extraction into w-relative clauses has a higher mean acceptability rate (4.3) than extraction into d-relative clauses (3.8). Again, it seems that LE into d-relative clauses is not absolutely unacceptable.

Summarizing our findings concerning LE into relative clauses, we found that w-relative clauses are at least as good as LE into embedded clauses. This suggests that the acceptability of LE is not affected by the landing site, similarly to the findings of the experiment in Section 4.1. On the basis of these findings, one would expect that the acceptability of long relativization would not be affected by the form of the relative pronoun. This prediction was not born out in the experiments comparing LE into w- and d-relative pronouns. We discuss this finding and its relevance for the findings of the main experiment in the discussion section. Finally, the overall acceptability scores suggest that LE enjoys a medial status, not being categorically accepted or rejected. Therefore, the acceptability of LE deserves more scrutiny.

4.3 LE in Embedded Clauses: Summary

To sum up, the experimental results reviewed in this section suggests that LE into embedded clauses is not categorically excluded in Modern German. On the contrary, all structures we tested were of medium acceptability. Of course, all instances of LE into an embedded clause were less acceptable than their control conditions. Interestingly, LE into relative clauses, which is supposed to be absolutely excluded from Modern German (e.g., Andersson & Kvam 1984; Lühr 1988), survived the tests. Another interesting finding obtained for relative clauses is a difference depending on the lexical type of the relative pronoun – w-relatives were judged better than d-relatives.

5 Experiment: Comparison Among Different LE Constructions

As mentioned above, LE into relative clauses has been claimed to be ungrammatical in Modern German (Andersson & Kvam 1984; Lühr 1988), contrary to our findings. Andersson & Kvam (1984) found that sentences with LE into a relative clause were most of the times rejected as ungrammatical. Andersson & Kvam (1984) had around 40 informants, but the sample of their sentences was very small. Consequently, the investigation of LE into relative clauses necessitates a more robust empirical basis. The difference between LE into d- and LE into w- relative clauses reported above enhances this necessity. In contrast to LE into relative clauses, LE into comparative clauses is supposed to be acceptable. However, these structures have not been compared to each other within one and the same experiment, to the best

of our knowledge. We therefore ran a new experiment comparing the four different types of LE.

In the present experiment, we compare four structures with LE, which differed with respect to type of the clause containing the landing site: comparative clauses (23a), embedded wh-questions (23b), relative clauses (23c), and declarative main clauses with long topicalization (23d). Each LE sentence was preceded by a context sentence. The context sentence was always the same for all four structures.

(23) Context: Lara macht ihre Ausbildung in einer Werkstatt und ist heute für die Reparaturen zuständig.

'Lara is trained in a workshop and is responsible for the repairs today.'

a. Comparative clause

Sie hat mehr Geräte repariert, als der Chef gesagt hat, dass sie reparieren soll.
she has more devices repaired than the director said has that she repair should

'She repaired more devices than the director said, that she should repair.'

b. Embedded w-question

Mich interessiert, wie viele Geräte der Chef gesagt hat, dass sie reparieren soll.
me interests how many devices the director said has that she repair should

'I want to know how many devices the director said that she has to repair.'

c. Relative clause

Sie hat wieder nicht die Geräte repariert, die der Chef gesagt hat, dass sie reparieren soll.
she has again not the devices repaired that the director said has that she repair should

'Once again, she failed to repair the devices which the director said that she should repair.'

d. Declarative main clause (long topicalization)

Sogar ein paar total alte Geräte hat der Chef gesagt, dass sie reparieren soll.
even one pair completely old devices has the director said that she repair should

'Even some totally old devices the director said that she should repair.'

On the basis of the claims of Andersson & Kvam (1984), we expect that comparative clauses have a higher acceptability rate than (at least) relative clauses. With regard to declarative main clauses (long topicalizations), Kiziak (2010) found a very low degree of acceptability. Since we use a context in our experiment and topicalization is therefore contextually licensed, acceptability should improve. Finally, we expect that LE into relative clauses will be (at least numerically) more acceptable than LE into embedded wh-questions, replicating the results of our previous experiment.

5.1 Method

Participants

24 students from the University of Frankfurt completed a questionnaire for course credit. All participants were native speakers of German and naïve with respect to the purpose of the experiment. 20 participants were from the state of Hesse. The remaining four participants were from the following four states: Bavaria, Baden-Wuerttemberg, Rhineland-Palatinate, and Mecklenburg-Hither Pomerania.

Materials

We constructed 20 experimental items, each appearing in the four conditions shown in example (23). Each item consisted of an introductory sentence followed by a target sentence with LE. The main purpose of the context sentence was to provide an appropriate context for sentences with long topicalization. For reasons of parity, the same introductory sentence also preceded the target sentence in the remaining conditions. The introductory sentence always contained a proper name referring to a character that was taken up in the target sentence using personal pronouns. The introductory sentences varied with regard to their syntactic structure and always provided some background information concerning the referent introduced by the proper name.

In each target sentence, the most deeply embedded clause was a *dass*-clause from which an accusative object was extracted. The subject of this clause was always a pronoun referring to the character introduced in the context sentence. The *dass*-clause was identical for the four experimental conditions. In the comparative clause condition (23a), the target sentence was a declarative main clause containing a subject pronoun and an accusative object consisting of the comparative operator *mehr* ('more') and a noun, which also occurred in the target sentence of the three other conditions. The main clause was followed by a comparative clause containing verbs taking clausal objects. The *dass*-clause containing the extraction site was the object of this verb. Target sentences in the *wh*-question condition (23b) began with a main clause containing a verb subcategorizing an indirect question. The main clause was followed by an indirect question starting with a *wh*-phrase consisting of the *wh*-determiner *wie viele* ('how many') and the same noun used with *mehr* in the comparative condition. The embedded question contained the same verb with clausal object as the comparative condition, and the *dass*-clause was again the object of this verb. Target sentences in the relative clause condition (23c) were similar to the sentences in the comparative condition, with the following differences. First, in the main clause the comparative operator was replaced by the definite article. Second, the comparative clauses were replaced by relative clauses with a *d*-relative pronoun. Sentences in the declarative condition (23d), finally, started with a main clause which had the moved object phrase in clause-initial position. This main clause contained the lexical material used in the other conditions as embedded clause sandwiched between main clause and *dass*-clause. In the declarative clause condition, the *dass*-clause was thus directly dependent on the main clause.

The 20 experimental items were distributed over four lists according to a Latin Square design. Each list contained only one version of each item and an equal number of items in each of the four conditions. The experimental sentences were interspersed into a list of 52 filler items. All filler items consisted of two sentences.

They were either fully acceptable or unacceptable in one of two ways. In one kind of unacceptable filler item, the two sentences were coherently related to each other, but one of them contained a grammatical error. In the other kind of unacceptable filler item, the two sentences were acceptable considered in isolation, but there was no coherent relation between them. The rationale for the fillers was that participants should be encouraged to pay attention both to the individual sentences but also to the connection between them.

Procedure

Four questionnaires were constructed on the basis of the four experimental lists. The experimental stimuli were randomized differently for each of the four questionnaires. Participants completed the questionnaires as part of a class session. They were asked to judge the acceptability of each item on the questionnaire by marking one of the numbers 1 to 7 printed beneath each sentence. A short instruction on the first page of the questionnaire told participants that 1 meant “totally unacceptable” and 7 meant “totally acceptable”, where acceptability could concern the two sentences of each item alone as well as the relationship between the two sentences. The instruction did not contain any example sentences. Participants needed about 15–20 minutes to complete the questionnaire.

5.2 Results

Table 1: Mean acceptability scores for the four constructions investigated in the present experiment. The standard error by participants is given in parentheses

| Comparative clause | Wh-question | Relative clause | Declarative clause (long topicalization) |
|--------------------|-------------|-----------------|--|
| 3.4 (0.36) | 2.8 (0.28) | 2.4 (0.26) | 2.2 (0.26) |

Table 1 presents the mean acceptability ratings for each of the four structures investigated in the present experiment. Comparative clauses received the highest ratings, followed by wh-questions and relative clauses. Declarative main clauses (long topicalizations) received the lowest ratings. In order to determine which of the mean ratings differed significantly from each other, we analyzed the data by means of mixed-effect modeling using the R statistics software version 3.1.1 (R Core Team 2016) and the R package lme4 (Bates et al. 2015). We used forward difference coding for testing the differences between adjacent factor levels. Contrast 1 tests the difference between comparative clauses and wh-questions. Contrast 2 tests the difference between wh-questions and relative clauses. Contrast 3 tests the difference between relative clauses and declarative main clauses (long topicalization). Following the advice given in Barr et al. (2013), we fitted a model including the full statistical model in the random effects for both participants and items.

Table 2 reports the full model summary as well as likelihood ratio tests (LRTs), which assess the contribution of single contrasts. Contrast 1 and thus the difference between comparative clauses (mean rating = 3.4) and wh-questions (mean rating = 2.8) is significant ($\chi^2 = 5.4728$, $p < .05$). Contrast 2 and thus the difference between wh-questions (mean rating = 2.8) and relative clauses (mean rating = 2.4) is also

significant ($\chi^2 = 5.164$, $p < .05$). Contrast 3, however, failed to reach significance ($\chi^2 = 2.5494$, $p = .11$). The ratings for relative clauses (mean rating = 2.4) and declarative main clauses (with long topicalizations) (mean rating = 2.2) thus do not differ reliably from each other.

Table 2: Results of mixed-effect model

| | Estimate | Std. Error | t value | p(LRT) |
|------------|----------|------------|---------|--------|
| Intercept | 2.7010 | 0.2660 | 10.155 | |
| Contrast 1 | 0.6577 | 0.2485 | 2.647 | < .05 |
| Contrast 2 | 0.3564 | 0.1822 | 1.956 | < .05 |
| Contrast 3 | 0.2234 | 0.1660 | 1.345 | .11 |

To get an impression of the individual variation, [Fig. 1](#) plots the individual results for the 20 Hessian participants. The grand mean for this group was 2.7, with a range from 1.0 to 4.8. [Fig. 1](#) reveals several noteworthy findings. First, while there are some participants who consistently assigned the lowest or almost lowest rating to each structure, the majority of participants have mean acceptability ratings of medium magnitude. This finding refutes claims to the effect that mean acceptability scores of medium size result from averaging across two groups of subjects, one with low and one with high mean ratings. Rather, at least in our data, the medium acceptability mean reflects the fact that most participants assigned a medium acceptability score. Second, comparative clauses are the only sentence type for which a subset of participants has high acceptability ratings. Thus, LE in comparative clauses seems to be completely acceptable for some speakers of our population, whereas all other structures are degraded at least to a certain extent. Third, [Fig. 1](#) also shows a large amount of individual variation, even though the regional background is held constant. Of course, given the coarse-grained classification in terms of German states, we cannot exclude that the variation seen in [Fig. 1](#) is due to more local regional variation. However, we consider it unlikely that the observed variation can be accounted for in this way. It is a task for future research to uncover the relevant factors responsible for this kind of individual variation. Much less individual variation was seen for the filler items. The majority of filler items consisted of two simple main clauses that were linked by referential continuity. Such items received mean ratings of about 5.5 on the 1-to-7 scale.

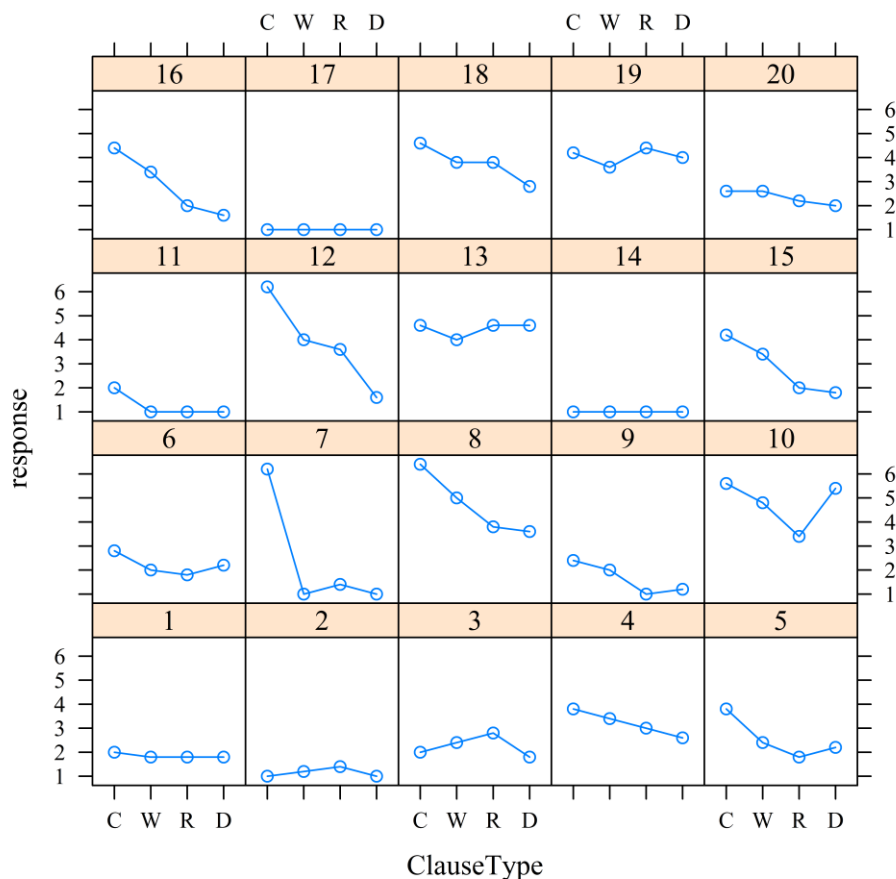


Fig. 1: Individual mean ratings for each of four structures for each of 20 Hessian participants

5.3 Summary

Comparative clauses were accepted to a higher degree than all other structures, as Andersson & Kvam (1984) suggested. Declarative main clauses (long topicalization) was the worst structure, despite the fact that they were provided with a context. This finding, which replicates the findings of Kiziak (2010), suggests that long topicalization per se is problematic. W-questions were judged better than relative clauses. This is in accordance with what we found in our previous experiments, since all relative clauses contained a d-pronoun, and such relative clauses were found to be less acceptable than relative clauses with a w-pronoun. Relative clauses and topicalization did not differ.

6 Discussion

The results presented in this paper show that LE in German is accepted to some degree in more syntactic contexts than assumed in much of the literature. LE into embedded questions and w-relative clauses is as acceptable as LE into main clause questions. LE into a comparative clause is even more acceptable as these three structures, whereas LE into a d-relative clause and a declarative main clause (long topicalization) received the lowest acceptability ratings.

An obvious question raised by the results presented in this paper concerns the reason for the variability in acceptability among the various constructions involving LE. Given the results discussed in this paper, the following ranking of structures involving LE can be established.

(23) Comparatives > wh-questions, w-relatives > d-relatives, declaratives

The question is how these differences come about, since we tested similar structures, i.e. structures in which an object NP is moved across the complementizer of a *dass*-sentence. At the moment, we can only make some tentative suggestions concerning some of the differences. In particular, we suspect that the differences are brought about by performance factors and factors related to information structure.

Consider first comparative clauses, which received the highest ratings of all constructions involving LE. LE in comparative clauses differs in two ways from LE in the other constructions under consideration. First, there is no alternative structure without LE available for comparative clauses, not even a colloquial one. Second, LE in comparative clauses does not involve movement of a phrase with special morphological marking (like wh-marking) to a designated specifier position like SpecCP. In other words, LE in comparative clauses does not involve movement of an active filler in the sense of Frazier (1987). Consequently, readers cannot predict a LE before they encounter the complementizer *dass*. That the comparative structures should be processed as LE becomes clear only when readers realize that the most deeply embedded *dass*-clause is missing the object required by the verb, that is, it contains a gap at the extraction site. Therefore, the advantage of the comparative clauses can be interpreted partially as an effect of local well-formedness. Taken together, these two properties may obscure the fact that these structures are instances of LE, making it less probable that participants with a negative attitude toward LE give such sentences a low rating too.

Consider next the finding that when targeted by LE, w-relative clauses and embedded questions had similar acceptability ratings, whereas d-relative clauses received lower ratings. The equivalence of w-relative clauses and embedded questions is probably least surprising given that the two structures are surface identical from the w-phrase onward. This is not the case for d-relative clauses, which differ from the other two structures precisely in that they contain a d-item instead of a w-item. As a consequence, the surface configuration of LE with d-relative clauses is much less frequent – if it occurs at all – than the surface configuration of LE with w-relative clauses and embedded questions. In addition, a subsequence like *was du glaubst* ('what you believe') does not only occur in sentences with LE, but also in sentences with partial movement, a construction that participants may be more acquainted with. The somewhat lower acceptability scores for d-relatives in comparison to w-relatives and embedded questions could thus simply be a consequence of the lower familiarity with the former than the latter.

A remarkable finding is also the rather poor acceptability of sentences with long topicalization. Our experiment replicated the findings reported in Kiziak (2010) despite the presentation of a context, which should make topicalization more natural. This finding suggests an approach to LE that is not purely syntactic. Specifically, if long A'-movement were the only factor affecting acceptability, we would expect the same acceptability scores across sentences. Since long topicalization is more often rejected as other structures with LE, we must assume that not syntax but rather

pragmatics is at play. The fact that the existence of a context did not improve the acceptability in comparison to the experiments of Kiziak (2010) makes the picture more complicated. We cannot exclude, of course, that our context was not appropriate to motivate topicalization.

On average, all of the structures received acceptability ratings that show that they are neither accepted without hesitation nor thoroughly rejected. As shown by the individual results in Fig. 1, the somewhat marginal status of sentences with LE is not an artifact of averaging across a group of participants. Our experimental findings concur with our corpus observations that authentic written examples of LE in German are either from older stages of German or from temporary colloquial German as found in informal internet bulletin boards, but rarely from websites adhering to a formal style (e.g., nationwide newspapers). In sum, we hypothesize that LE is an integral part of the grammar of German which has acquired a somewhat marginal status due to normative pressure.

References

- Andersson, S.-G. & S. Kvam (1984) *Satzverschränkung im heutigen Deutsch*. Narr, Tübingen.
- Bader, M. & V. Koukouloti (in preparation) Long extraction in German: banned but still alive. Manuscript in preparation, University of Frankfurt.
- Baroni, M., S. Bernardini, A. Ferraresi, & E. Zanchetta (2009) The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation Journal*, 23(3): 209-226.
- Barr, D. J., R. Levy, C. Scheepers, & H. J. Tily (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68: 255-278.
- Bates, D., M. Mächler, B. Bolker, & S. Walker (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1): 1-48.
- Brandt, P. & E. Fuß (2014) Most questionable pronouns: Variation between das- vs. was-relatives in German. *Linguistische Berichte*, 239: 297-329.
- Chomsky, N. (1977) On wh-movement. In P. Culicover, T. Wasow, & A. Akamajian, eds., *Formal Syntax*. Academic Press, New York: 71-132.
- Fanselow, G. (1987) *Konfigurationsalität. Untersuchungen zur Universalgrammatik am Beispiel des Deutschen*. Narr, Tübingen.
- Fanselow, G. & S. Frisch (2006) Effects of processing difficulty on judgments of acceptability. In G. Fanselow, C. Féry, R. Vogel, & M. Schlesewsky, eds., *Gradience in grammar: Generative perspectives*. Oxford University Press, Oxford/New York: 91-316.
- Featherston, S. (2005) That-trace in German. *Lingua*, 115(9): 1277-1302.
- Frazier, L. (1987) Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5: 519-559.

- Grewendorf, G. (1995) Syntaktische Skizzen. In J. Jacobs, A. von Stechow, W. Sternefeld, & T. Vennemann, eds., *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*. De Gruyter, Berlin/New York: 1288-1319.
- Hoeksema, J. & A. Schippers (2012) Diachronic changes in long-distance dependencies. In A. v. Kemenade, ed., *Historical Linguistics 2009: Selected Papers from the 19th International Conference on Historical Linguistics, Nijmegen, 10-14 August 2009*. John Benjamins, Amsterdam/Philadelphia: 155-170.
- Kiziak, T. (2010) *Extraction Asymmetries: Experimental Evidence from German*. John Benjamins, Amsterdam/Philadelphia.
- Lühr, R. (1988) Satzverschränkung im heutigen Deutsch. *Groninger Arbeiten zur Germanistischen Linguistik*, 28: 73-87.
- Müller, G. & J. Sabel (1989) Aufgebrochene Barrieren. *Frankfurter Linguistische Forschungen*, 6: 22-40.
- Paul, H. (1916/19) *Deutsche Grammatik*. Max Niemeyer, Halle.
- R Core Team (2016) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org/> (last access March 20, 2017).