Word order processing in German as a foreing language

Teodor Petrič

2022-04-25

### Abstract

In an experiment on processing effort in sentence comprehension, preference hypotheses on word order and relative clause position were tested in German as a second or foreign language. The experiment involved 62 Slovenian students of German studies in Maribor as participants. The linguistic material consisted of 16 utterances, each with a main clause and a relative clause and each of them subject-initial or object-initial. The relative clause could occur either after the first or second noun phrase. The procedure used was the self-paced reading paradigm. After each test unit, the participants were asked a comprehension question, either about the subject of the main clause (*who hit?*) or about the subject of the relative clause (*who saw?*). The experiment confirmed the hypotheses that response accuracy to a comprehension question was higher when a clause was subject-initial and when the relative clause occurred after the first noun phrase of the main clause. The experiment did not confirm the assumed interaction between word order and position of the relative clause.

### Keywords

sentence comprehension, word order, subject-initial preference, noun phrase accessibility hierarchy, relative clause, German, Slovenian, self-paced reading paradigm, second language acquisition, naturalness theory.

## 1 Introduction

Previous research findings (overview in ([Weyerts et al.](#ref-weyerts2002word))) suggest that sentences are processed incrementally. The comprehension of a sentence is influenced by a number of linguistic and extra-linguistic factors. A significant grammatical influence is word order. For example, the preference for the sequence *subject (S) before object (O)* is rated to be a typologically unmarked or more natural option, being prevalent in about 96% out of more than 400 languages studied ([Tomlin](#ref-tomlin2014basic)). In psycholinguistic studies with native speakers (overview in ([Kaan](#ref-kaan1997processing))), it has also been demonstrated that sentences with SO constituent order are easier to understand than those with reverse order (OS).

It is known from studies with native-speaking subjects that certain positional patterns are preferred. This also raises the question of how the *noun phrase accessibility hierarchy* ([Keenan and Comrie](#ref-keenan1977noun)) and the position of an attributive relative clause influence the comprehension of a sentence. Our online experiment with learners of German as a second or foreign language aims to test whether certain positions of the subject, the object and the respective relative clause facilitate or impair sentence comprehension. The design of our experiment is based on an experiment originally conducted by ([Junker](#ref-junker2004verflixte)) with German native speakers.

### 1.1 Word order

Several psycholinguistic studies have investigated the influence of constituent order on the comprehension of sentences (overview in ([Weyerts et al.](#ref-weyerts2002word))). While in English the preverbal NP is typically interpreted as the subject, even if morphological markers (e.g. congruency markers) suggest a correction of the perspective, in morphologically richer languages (e.g. German, Slovenian) one also resorts more often to morphological and semantic properties of a noun to infer, for example, the subject of a sentence. The preference for the sequence subject before object (SO order) is proven in many languages. Response time differences in languages with flexible constituent order (German, Dutch, Russian, Serbian/Croatian, Finnish, Japanese) indicate that canonical SO order is easier to process than non-canonical (i.e. those with OS order).

Besides grammatical properties such as constituent order, discourse factors e.g. context, thematic progression ([Kaiser and Trueswell](#ref-kaiser2003putting)) also influence the processing speed of constituent sequences (SVO vs. SOV vs. OVS …). Thematic progression or the distribution of information influences the formal complexity of our sentences, for example the choice of a noun or pronoun. Often, given sentence units are realised as unstressed pronouns. Like other thematic elements, they usually show a tendency to appear near the beginning of a sentence. As a result, the otherwise preferred SO sequence is reversed.

### 1.2 Noun Phrase Accessibility Hierarchy

The Noun Phrase Accessibility Hierarchy (NPAH) was introduced by ([Keenan and Comrie](#ref-keenan1977noun)), i.e. an implicature scale for the relativisability of syntactic relations ([Fox](#ref-fox1987noun): 856). Accordingly, the languages of the world follow the following implicature scale:

**subjects > direct objects > indirect objects > oblique objects > genitives > object complements**.

According to this, it should be possible in every language to modify a subject with a relative construction, but it should not be possible in all of them to modify other syntactic relations (e.g. the direct object) in this way. In the latter case, in some languages speaker resort to so-called primary strategies, i.e. that they choose, for example, passive verb constructions to transform the direct object into a subject and thus make it amenable to relativisability. ([Keenan and Comrie](#ref-keenan1977noun): 652)

The hierarchy of relativisability is supported by the varying frequency of relative clauses along the implicature scale. In English texts, relative clauses to subjects appear more frequently than to (direct) objects, the latter more frequently than obliques, etc. ([Keenan](#ref-keenan1975variation)) Another finding cited was the relationship between subject and object relative clauses (subject and object, respectively, are clause members of the relative clause in this case), which correlates with the complexity of (English) texts ([Keenan](#ref-keenan1975variation)): simpler texts showed a greater difference between subject and object relative clauses than more complex texts. Even in a language like English, where relative clauses are permissible for all syntactic relations on the implicature scale, statistically relevant differences are thus evident. German and Slovenian presumably have similar conditions to English.

In accordance with the Relativisability Hierarchy ([Keenan and Comrie](#ref-keenan1977noun)), it should be possible to predict the difficulty level and acquisition order of relative clauses in L2 acquisition. However, the generalisability of the implicature scale is limited by the fact that most studies on NPAH have been conducted with European languages. ([Ozeki and Shirai](#ref-ozeki2007does)) For example, an analysis of relative clauses of non-native Japanese speakers found that even learners with poorer language skills used object and oblique relations. This suggests that subject relations are not easier for L2 Japanese learners than object or oblique relations and that learners use different types of relative clauses in accordance with the animacy of the head noun. ([Ozeki and Shirai](#ref-ozeki2007does))

## 2 Word order in German and Slovenian

### 2.1 Word order in German

German is classified as a language with mixed constituent order ( and ) according to ([Dryer](#ref-wals-81)) or as an OV-language with an V-second property on top ([Abraham](#ref-abraham1995deutsche); [Haftka](#ref-haftka2019deutsch); [Haider and Szucsich](#ref-haider2018slavic)). In the topological sentence field model (cf. ([Abraham](#ref-abraham1995deutsche); [Haftka](#ref-haftka2019deutsch); [Haider and Szucsich](#ref-haider2018slavic))), which we choose for description, subject and object can occur either in the front field (i.e. before the finite verb form), in the middle field (i.e. between the finite and infinite verb forms) or, in the case of sentence-like realisation, also in the final field of a sentence (i.e. after the last verb form). In declarative main clauses such as (1) and (2) and in interrogative clauses with wh-elements (5), a sentence constituent (e.g. the subject or the object) is in the front field and the finite verb form is in the left-hand clause frame. In embedded clauses (3), the finite verb form appears (together with the infinite verb form) in the right-hand clause frame. In some constructions, e.g. in yes-no questions (4), conditional and imperative sentences, the finite verb form is always or often the initial element of the sentence. According to the sentence field model, the finite verb is assumed to be in the same position as in yes-no questions following an empty (or non-generated) front field. In contrast to most languages of the world ([Comrie and Kuteva](#ref-wals-122)), German uses relative pronouns to establish a reference to the antecedent in the main clause. Example (3) shows a relative clause to the subject *der neue Song* in the main clause, where the relative pronoun *den* is an accusative object in the relative clause, but not case-congruent with the subject in the main clause.

| Beispiel | Äußerung |
| --- | --- |
| 1 | Capital Bra hat einen neuen Song veröffentlicht. |
| 2 | Einen neuen Song hat Capital Bra veröffentlicht. |
| 3 | Der neue Song, den Capital Bra veröffentlicht hat, ... |
| 4 | Hat Capital Bra einen neuen Song veröffentlicht? |
| 5 | Wann hat Capital Bra einen neuen Song veröffentlicht? |

### 2.2 Word order in Slovenian

Slovenian, a morphologically rich language with largely fusional features, is often characterised as a language with flexible constituent order and dominant SVO order. But like other Slavic languages, Slovenian lacks many of the defining structural properties of typical SVO languages (e.g., English) and shares several properties with SOV languages so that alternative syntax taxonomies which include ambidirectional licensing of a verbal head should be discussed ([Haider and Szucsich](#ref-haider2018slavic)). In contrast to German, Slovenian also belongs to the pro-drop languages. Slovenian uses relative pronouns ([Toporišič](#ref-toporivsivc1992slovenska); [Greenberg](#ref-greenberg2006short)) which are also used as interrogative pronouns (e.g. *kateri*, “which one”), but even more frequent is the use of the uninflected (subjunctive) relative pronoun *ki* or its combination with a cliticised personal pronoun in the left part of the relative clause (e.g. *jo*, “her” as in *ženska, ki sem jo ljubil* > woman, Relpron was-I her loved > “the woman I (have) loved”). This construction is also common in colloquial Czech. ([Keenan and Comrie](#ref-keenan1977noun))

| Beispiel | Äußerung |
| --- | --- |
| 6 | Adele je ta teden objavila svoj novi album. Adele - Aux - this week - published - her new album |
| 7 | Svoj novi album je ta teden objavila Adele. Her new album - Aux - this week - published - Adele |
| 8 | Plošča, ki jo je Adele ta teden objavila, ... Record - Relpron - her (clitic) - Aux - Adele - this week - published |
| 9 | (Ali) je Adele ta teden objavila svoj novi album? (question particle) - Aux - Adele - this week - published - her new album |
| 10 | Naslednji teden pevka, ki je ta teden objavila svojo novi album, nastopa v Londonu. Next week - singer, - Relpron - Aux - this week - published - her new album, - performs - in London |
| 11 | Naslednji teden se pevka vrne iz Pariza. Next week - herself - singer - returns - from Paris |

The dominant verb-before-object order is observed both in the main clause (6) and in the embedded clause (8). In emphasis (7), the rhematic subject can also go to the end of the sentence and the object to the front. In yes-no questions (9), a question particle (*ali*) can appear. In the second position there is a clitic, namely the auxiliary verb *biti* (“to be”) or other sentence clitics (e.g. unstressed personal pronouns in different case or unstressed reflexive pronoun *se*). It is also possible to begin a sentence with a (pro)clitic or clitic cluster (on clitics in Slovenian and/or other Slavic languages cf. ([Franks and King](#ref-franks2000handbook); [Bošković](#ref-bovskovic2016second); [Golden and Sheppard](#ref-golden2000slovene); [Marušič](#ref-maruvsivc2009positioning))).

In an example with a relative clause (10), a topicalised phrase *naslednji teden* (“next week”) appears before the subject of the main clause *pevka* (“singer”). The finite verb form in the main clause *nastopati* (“to perform”) is not in the head position of the complementizer phrase (CP), but presumably ([Petrič](#ref-petric1998verbposition)) in the head position of the inflectional phrase (IP). In example (11), the cliticised reflexive pronoun *se* (“itself”) occurs in second position (attached to the complementizer phrase), the finite verb form *vrne* of the reflexive main verb *vrniti se* (“to return”) is not raised into the complementizer phrase (CP) and remains in the inflectional phrase (IP).

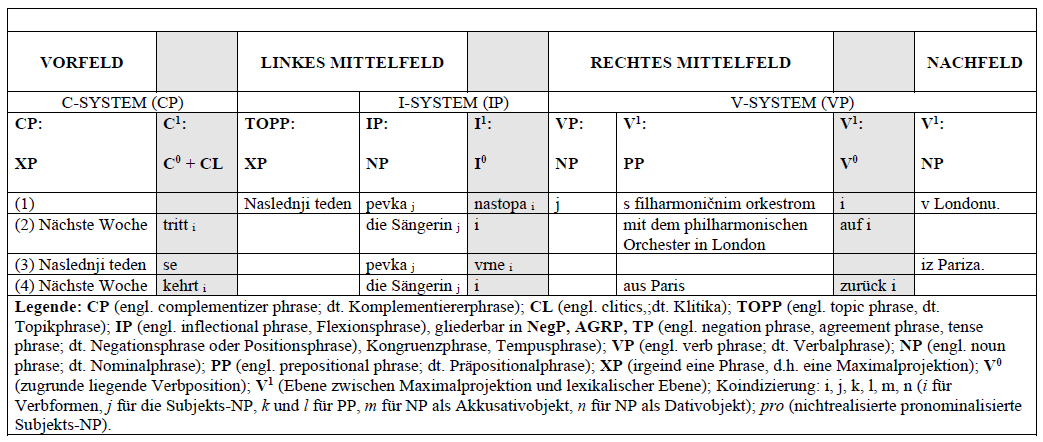


Table: Sentence fields in German and Slovenian ([Petrič](#ref-petric1998verbposition))

From a surface syntactical point of view, it remains to be noted that in Slovenian the constituent order is preferred in both main and subordinate clauses and that the object only ends up before the subject when it is topicalised or pronominalised (due to information-structure fitting). Relative clauses can be connected to various syntactic relations along the implicative relativisability scale, similar to German or English.

## 3 Theoretical background and hypotheses

Using the conceptual inventory of naturalness theory ([Dressler et al.](#ref-dressler1987leitmotifs)), we evaluate the serialisation types SVO and SOV as more natural in comparison to those in which the object appears before the subject (i.e.: OSV, OVS, VOS), since in the former the preferential subject-first principle dominates, whereas in the latter it does not. In languages with a dominant SO order, the agent of an action is realised before the object (which is produced or changed by the agent). This corresponds to the anthropocentric view in communication. This naturalness assessment is supported by results of cross-linguistic research ([Dryer](#ref-wals-81)), according to which most languages in the world have either a dominant SVO or a SOV sequence.

Based on the hierarchy of relativisability, we consider a relative clause to the subject of the main clause to be more natural than a relative clause to another syntactic relation of the main clause (e.g. to the direct object). For learners of German as a second or foreign language, an utterance should therefore be easier to understand if the relative clause refers to the subject of the main clause and not to the object.

Processing a relative clause that does not refer to the first noun but to one that occurs later in the clause has the disadvantage that the recipient has to remember several potential referents and that this places a greater burden on working memory. Therefore, we evaluate the position of the relative clause after the first nominal phrase (position 1) as a more natural option than after a second or later one (position 2).

Because of the expected preference for subject-initial sentences and the preference for relative clauses to the subject, we expect relative clauses after the subject in position 1 to be easier to understand than relative clauses after the subject or another syntactic relation in position 2.

Based on the assumption that less natural constructions require more cognitive effort than corresponding natural constructions, it is to be expected that the former require more processing time and / or additional memory resources than the latter. A Slovenian learner of German as a second or foreign language should thus understand German sentences with less natural OS sequence not as well than sentences with the more natural SO sequence. Experimentally, the greater cognitive effort should be reflected in increasing processing time for the presented test units and decreasing processing accuracy when answering the comprehension questions.

The operationalised hypotheses tested in our experiment can be formulated as follows:

() Comprehension questions about a subject-initial clause (SO-sequence) will be answered correctly more often than those after an object-initial clause (OS-sequence). This hypothesis invokes the typologically based preference of subject-initial sentences.

() Comprehension questions about a sentence with a relative clause in the first position (i.e. after the first noun in the utterance) are answered correctly more often than those with a relative clause in the second position (i.e. after the second noun in the utterance). This hypothesis is based on the assumption that when processing a relative clause that appears late, one has to remember several potentially relevant referents and that this places a greater burden on working memory.

() Comprehension questions on a subject-initial sentence with a relative clause in position 1 are answered correctly more often than those on subject- or object-initial sentences with a relative clause in position 2. The interaction between syntactic relation and position of the relative clause is expected due to the load on working memory, the preference for subject-initial sentences described above and the implicative relativisability scale for syntactic relations.

## 4 Experiment

A well-established psycholinguistic method for determining the processing effort of sentence comprehension is the self-paced reading paradigm ([Gibson](#ref-gibson1998linguistic)). According to the hypotheses formulated above, the author expects that the preferred SO order in German will require shorter response times and facilitate comprehension of sentence content than the non-preferred OS sequence. To check whether the predictions made above are relevant in second and foreign language acquisition, we measured the responses of our participants to main and relative clauses with SO and OS order using a self-paced reading task ([Just et al.](#ref-just1982paradigms); [Nerantzini et al.](#ref-nerantzini2022processing)).

### 4.1 Participants

In contrast to the experiment in ([Junker](#ref-junker2004verflixte)), the experiment was not conducted with German first language learners, but with 62 Slovenian participants who had acquired German as a foreign language and were studying German at the University of Maribor. The participants had advanced German language skills, in most cases at level B2 according to the European Referential Framework of languages, in a few cases even above. Therefore, we expected that the cognitive processing in our participants would not differ in principle from the processing in native speakers.

### 4.2 Linguistic material

Two experimental versions were constructed: a longer one with 16 different complex sentences and a shorter one with a single complex sentence as the basis for manipulations of word order and relative clause position. As we were unable to recruit enough participants for the longer experiment (which consisted of 16 versions of each of the 16 complex sentences), we only report details of the shorter experiment. In our experiment, test units were formed similar to those constructed in the study of ([Junker](#ref-junker2004verflixte)). The participants were gradually given utterances to read that consisted of a main clause with an embedded relative clause and contained a total of three nouns and two verbs (one in the main clause, the other in the relative clause). Both the main and relative clauses could begin with either a subject or an object. The position of the relative clause was also varied (either after the first or the second noun of the main clause). The nouns and relative pronouns appeared either in the nominative or accusative singular. The subject-object order in the main and relative clauses could be congruent or not. After an utterance was read, the participants were asked a comprehension question using active voice.

#### 4.2.1 Variation of constituent and clause position

The experiment followed a 2x2x4 design with 16 test utterances, i.e. with two sentence types (main clause, relative clause), two positions of the relative clause (position after noun or noun) and four possible (congruent or non-congruent) subject-object sequences in main and relative clause (SO + SO, OS + SO, SO + OS, OS + OS). Each test unit consisted of 12 word segments. In all 16 utterances, the nouns *Mann, Clown, Anwalt* (“man, clown, lawyer”) and the relative pronoun (*der, den*, “the”) appeared in the nominative or accusative singular. In the main clause, the past participle of the verb *schlagen* (“to beat”) was used, whereas in the relative clause, that of the verb *sehen* (“to see”) was used. The following examples illustrate the experimental conditions.

| Item | Testeinheiten | Hauptsatz | Relativsatz | Position |
| --- | --- | --- | --- | --- |
| 1 | Zweifellos hat der Mann der den Clown gesehen hat den Anwalt geschlagen | SO | SO | 1 |
| 2 | Zweifellos hat den Mann der den Clown gesehen hat der Anwalt geschlagen | OS | SO | 1 |
| 3 | Zweifellos hat der Mann den der Clown gesehen hat den Anwalt geschlagen | SO | OS | 1 |
| 4 | Zweifellos hat den Mann den der Clown gesehen hat der Anwalt geschlagen | OS | OS | 1 |
| 5 | Zweifellos hat der Mann der den Clown gesehen hat den Anwalt geschlagen | SO | SO | 1 |
| 6 | Zweifellos hat den Mann der den Clown gesehen hat der Anwalt geschlagen | OS | SO | 1 |
| 7 | Zweifellos hat der Mann den der Clown gesehen hat den Anwalt geschlagen | SO | OS | 1 |
| 8 | Zweifellos hat den Mann den der Clown gesehen hat der Anwalt geschlagen | OS | OS | 1 |
| 9 | Zweifellos hat der Mann den Anwalt der den Clown gesehen hat geschlagen | SO | SO | 2 |
| 10 | Zweifellos hat den Mann der Anwalt der den Clown gesehen hat geschlagen | OS | SO | 2 |
| 11 | Zweifellos hat der Mann den Anwalt den der Clown gesehen hat geschlagen | SO | OS | 2 |
| 12 | Zweifellos hat den Mann der Anwalt den der Clown gesehen hat geschlagen | OS | OS | 2 |
| 13 | Zweifellos hat der Mann den Anwalt der den Clown gesehen hat geschlagen | SO | SO | 2 |
| 14 | Zweifellos hat den Mann der Anwalt der den Clown gesehen hat geschlagen | OS | SO | 2 |
| 15 | Zweifellos hat der Mann den Anwalt den der Clown gesehen hat geschlagen | SO | OS | 2 |
| 16 | Zweifellos hat den Mann der Anwalt den der Clown gesehen hat geschlagen | OS | OS | 2 |

#### 4.2.2 Comprehension questions

After each test unit, the participants were asked a comprehension question in active voice, either one referring to the main clause (with the past participle of the verb *schlagen*) or to the relative clause (with the past participle of the verb *sehen*). In contrast to ([Junker](#ref-junker2004verflixte)), no comprehension questions were used in the passive voice in our experiment.

| Verständnisfrage | Referenz |
| --- | --- |
| Wer hat geschlagen? | Bezug auf den Hauptsatz |
| Wer hat gesehen? | Bezug auf den Relativsatz |

### 4.3 Procedure

Before the experiment, the participants were shown a description of the task on the computer screen and instructions for pressing the corresponding keys on the computer keyboard. In a short practice phase, the participants could familiarise themselves with the procedure and task.

The subjects were all unfamiliar with the aim of the experiment and participated voluntarily. Participants performed a self-paced reading task in which they had to press the key given in the instruction to display consecutive sentence segments on the computer screen. The presentation of each test unit began with a structural grid of the words (not yet visible) in the centre of the screen, with each line corresponding to a word. After pressing a key (*G* key), the first word (always the sentence adverb *zweifellos*, “unquestionably”) appeared on the left initial edge of the grid. After pressing the next key (space bar, left or right shift key), the word disappeared from the first position in the sentence grid and after about 200 milliseconds, the next word appeared at the second position in the sentence grid, namely the auxiliary verb *haben* (“to have”) in the third person singular passive indicative. Each utterance began with a main clause. After each keystroke, the participants saw a new word at the next position in the sentence grid and the previous one disappeared. To facilitate the task, commas in the sentence grid marked the expected position of the relative clause.

Each test unit was thus presented as a twelve-part structural grid, whereby the participants only ever saw one word on the screen, which was replaced by a word at the next position in the sentence grid after pressing a button. The maximum reading time for a sentence segment was limited to nine seconds. After that, the next segment was automatically shown.

| Item | Seg1 | Seg2 | Seg3 | Seg4 | Seg5 | Seg6 | Seg7 | Seg8 | Seg9 | Seg10 | Seg11 | Seg12 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | Zweifellos | hat | der | Mann | der | den | Clown | gesehen | hat | den | Anwalt | geschlagen |
| 2 | Zweifellos | hat | den | Mann | der | den | Clown | gesehen | hat | der | Anwalt | geschlagen |
| 3 | Zweifellos | hat | der | Mann | den | der | Clown | gesehen | hat | den | Anwalt | geschlagen |
| 4 | Zweifellos | hat | den | Mann | den | der | Clown | gesehen | hat | der | Anwalt | geschlagen |
| 5 | Zweifellos | hat | der | Mann | der | den | Clown | gesehen | hat | den | Anwalt | geschlagen |
| 6 | Zweifellos | hat | den | Mann | der | den | Clown | gesehen | hat | der | Anwalt | geschlagen |
| 7 | Zweifellos | hat | der | Mann | den | der | Clown | gesehen | hat | den | Anwalt | geschlagen |
| 8 | Zweifellos | hat | den | Mann | den | der | Clown | gesehen | hat | der | Anwalt | geschlagen |
| 9 | Zweifellos | hat | der | Mann | den | Anwalt | der | den | Clown | gesehen | hat | geschlagen |
| 10 | Zweifellos | hat | den | Mann | der | Anwalt | der | den | Clown | gesehen | hat | geschlagen |
| 11 | Zweifellos | hat | der | Mann | den | Anwalt | den | der | Clown | gesehen | hat | geschlagen |
| 12 | Zweifellos | hat | den | Mann | der | Anwalt | den | der | Clown | gesehen | hat | geschlagen |
| 13 | Zweifellos | hat | der | Mann | den | Anwalt | der | den | Clown | gesehen | hat | geschlagen |
| 14 | Zweifellos | hat | den | Mann | der | Anwalt | der | den | Clown | gesehen | hat | geschlagen |
| 15 | Zweifellos | hat | der | Mann | den | Anwalt | den | der | Clown | gesehen | hat | geschlagen |
| 16 | Zweifellos | hat | den | Mann | der | Anwalt | den | der | Clown | gesehen | hat | geschlagen |

In addition, to ensure that the participants had actually read and understood the sentence segments shown step by step on the screen, a question was asked after a test unit that referred either to a noun in the main clause (*Wer hat geschlagen?*, “who hit?”) or to one in the relative clause (*Wer hat gesehen?*, “who saw?”). Participants answered the comprehension question by pressing one of three keys on the computer keyboard: either the left shift key (for the phrase *der Mann*, “the man”, marked in red), the space bar (for the phrase *der Clown*, “the clown”, marked in purple) or the right shift key (for the phrase *der Anwalt*, “the lawyer”, marked in green). The time for answering the comprehension question was again limited to nine seconds. If a key was not pressed, the structural pattern of the next test unit appeared on the screen. The test units were pseudo-randomised similar to ([Junker](#ref-junker2004verflixte)) to ensure that the participants were not shown the test units in an identical order.

In the instruction phase, participants were asked to solve the task as quickly and as correctly as possible. Self-paced reading time was determined and measured as the time interval between the button press for one segment to appear and the button press for the next segment to appear. The response options were time-limited.

The presentation of the linguistic stimuli, the randomisation of the test units and the measurement of response time and accuracy were carried out using the psycholinguistic software *DmDx* ([Forster and Forster](#ref-forster2003dmdx)) and then analysed using the statistical software R ([Wickham et al.](#ref-wickham2019tidyverse); [Bates et al.](#ref-bates2015lmer); [Bartoń](#ref-barton2020inference); [Gohel](#ref-gohel2021flex); [Shenoy](#ref-shenoy2021grafify)).

## 5 Results

### 5.1 Descriptive Statistics

The comprehension questions were answered by our participants after an average of 3921 milliseconds (ms). The percentage of correct answers to the comprehension questions was about 61%, though slightly higher (but not significantly) for questions about the subject of a relative clause (62%). The average response accuracy (i.e. the proportion of correct answers to comprehension questions) was therefore significantly higher than the guessing probability of 33% correct answers. Correct answers were given more promptly (on average after 3183 ms) than incorrect answers (on average after 5100 ms).

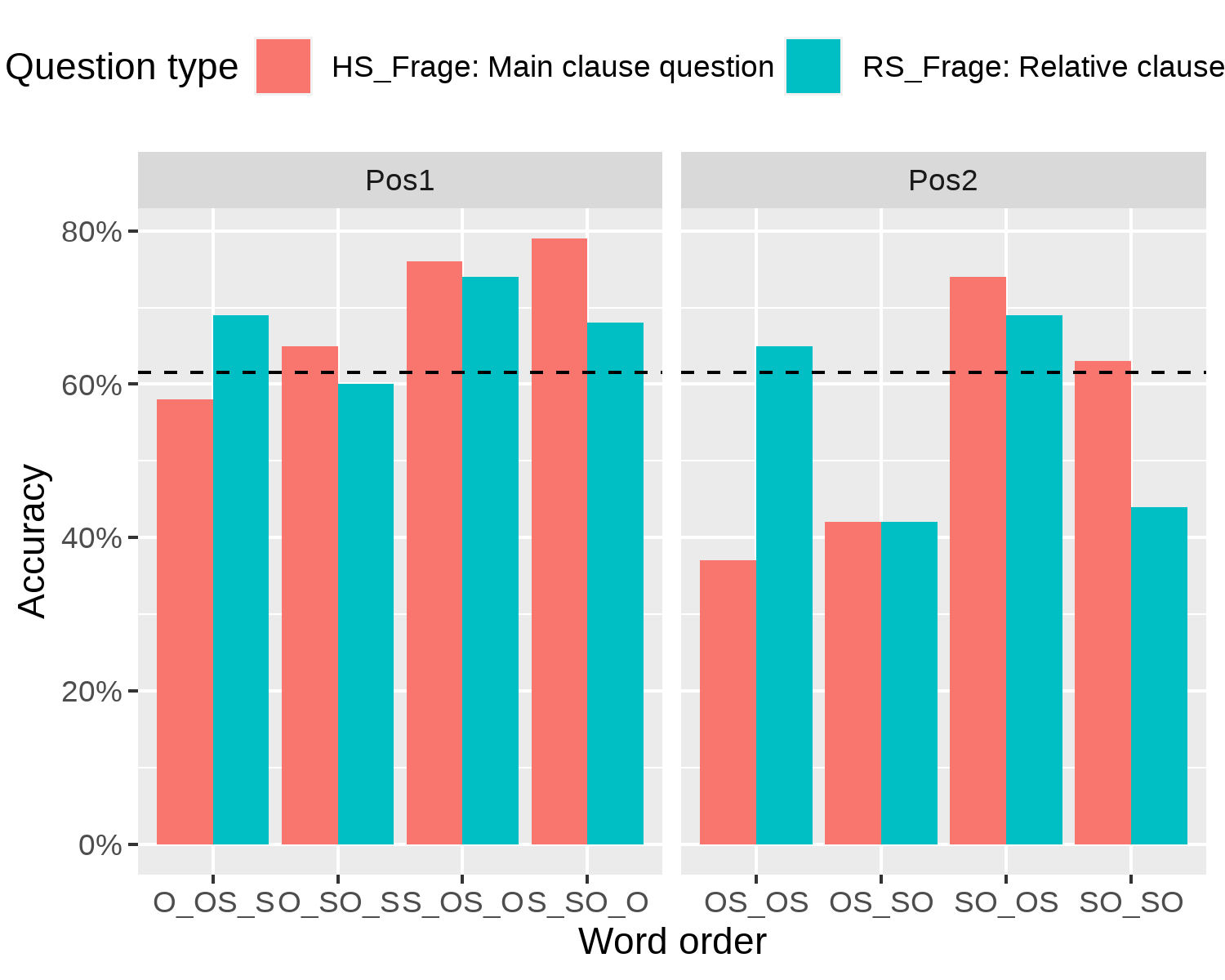
A question about the subject in the main clause (HS question [HS\_Frage]: *Who hit?*) was answered more slowly on average (about 4044 ms) than a question about the subject in the relative clause (RS question [RS\_Frage]: *Who saw?*, 3798 ms). However, the response time difference was only marginally significant (t(973.05) = 1.70; p = 0.09).

The response times for correctly answered HS questions were slightly shorter (by 319 ms) than for correctly answered RS questions (3343 ms). The difference was significant: t(605.05 = 2.67; p = 0.008. Incorrect answers to HS questions, on the other hand, were strongly delayed (on average after 5686 ms) compared to incorrect answers to RS questions (4519 ms). This response time difference was also statistically significant: t(379.18 = 4.21; p < 0.001).

The percentage of correctly answered questions was highest when the main clause was subject-initial and the relative clause in position 1 (74 - 79%). The percentage of correctly answered questions was lowest for object-initial main clauses and relative clause in position 2 (37 - 42%). The word order in the relative clauses had no noticeable effect on the proportion of correctly answered questions, but the position of the relative clause did. In position 2, mostly lower response accuracy was observed (the lowest at 37 - 44%).

| Fragetyp | HS-Abfolge | RS-Abfolge | Position | Anteil korrekter Antworten | | Reaktionszeit bei Frage [ms] | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | (Mittel) | (Stdev.) | (Mittel) | (Stdev.) |
| HS\_Frage | SO | SO | Pos1 | 0,79 | 0,41 | 3.689 | 2.325 |
| HS\_Frage | SO | OS | Pos1 | 0,76 | 0,43 | 3.896 | 2.107 |
| HS\_Frage | SO | OS | Pos2 | 0,74 | 0,44 | 4.167 | 2.513 |
| RS\_Frage | SO | OS | Pos1 | 0,74 | 0,44 | 3.188 | 1.649 |
| RS\_Frage | OS | OS | Pos1 | 0,69 | 0,46 | 3.308 | 1.653 |
| RS\_Frage | SO | OS | Pos2 | 0,69 | 0,46 | 3.756 | 2.045 |
| RS\_Frage | SO | SO | Pos1 | 0,68 | 0,47 | 4.079 | 2.254 |
| HS\_Frage | OS | SO | Pos1 | 0,65 | 0,48 | 3.941 | 2.298 |
| RS\_Frage | OS | OS | Pos2 | 0,65 | 0,48 | 3.485 | 2.080 |
| HS\_Frage | SO | SO | Pos2 | 0,63 | 0,49 | 3.873 | 2.556 |
| RS\_Frage | OS | SO | Pos1 | 0,60 | 0,49 | 4.072 | 2.303 |
| HS\_Frage | OS | OS | Pos1 | 0,58 | 0,50 | 3.984 | 2.437 |
| RS\_Frage | SO | SO | Pos2 | 0,44 | 0,50 | 4.213 | 2.292 |
| HS\_Frage | OS | SO | Pos2 | 0,42 | 0,50 | 4.276 | 2.440 |
| RS\_Frage | OS | SO | Pos2 | 0,42 | 0,50 | 4.283 | 2.334 |
| HS\_Frage | OS | OS | Pos2 | 0,37 | 0,49 | 4.524 | 2.653 |

The following picture shows that the proportion of correct answers to HS questions (which focus on the subject of the main clause) was greater for subject-initial main clauses (SO order) than for object-initial main clauses. This effect diminished with RS questions, especially with relative clauses in position 2.

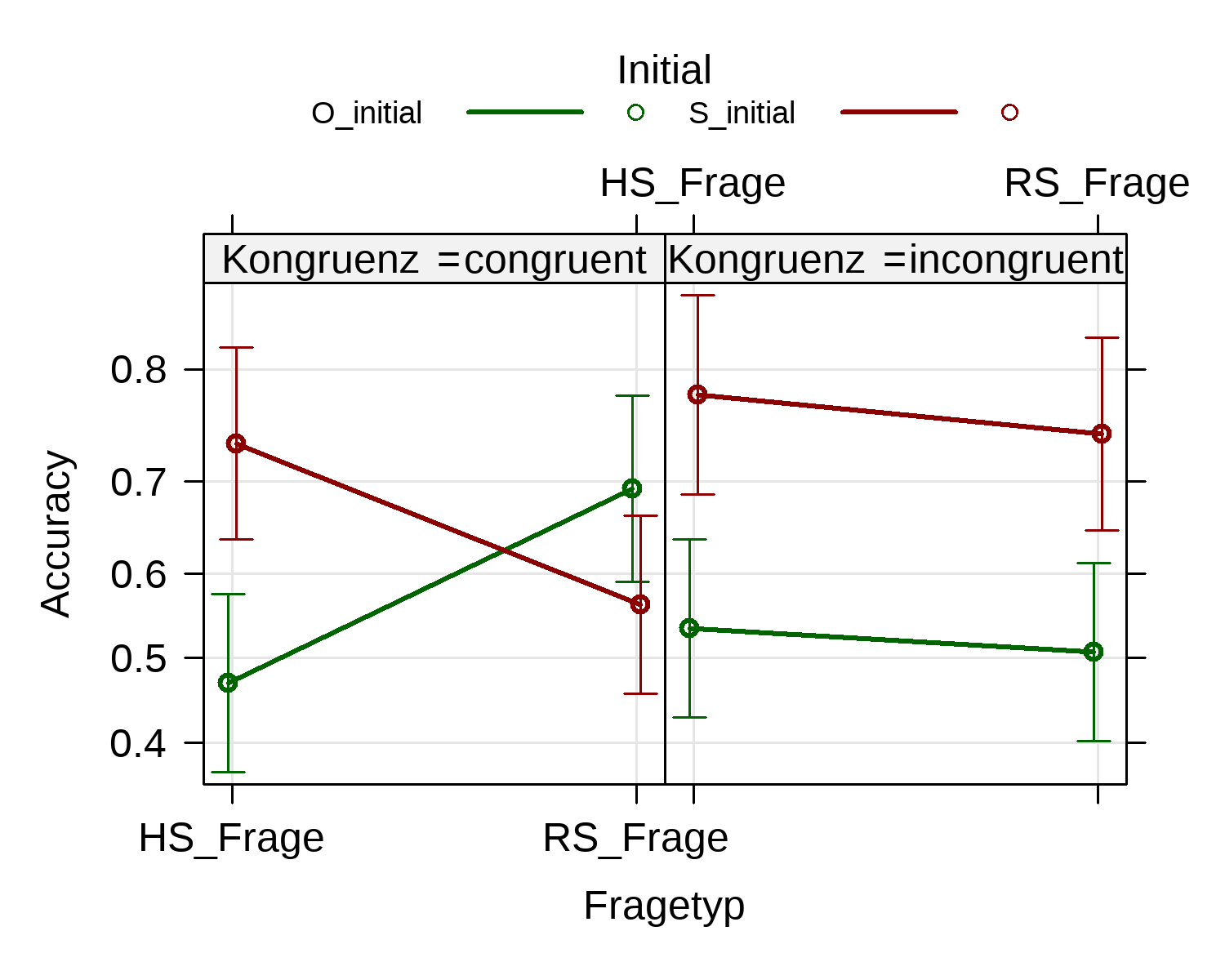
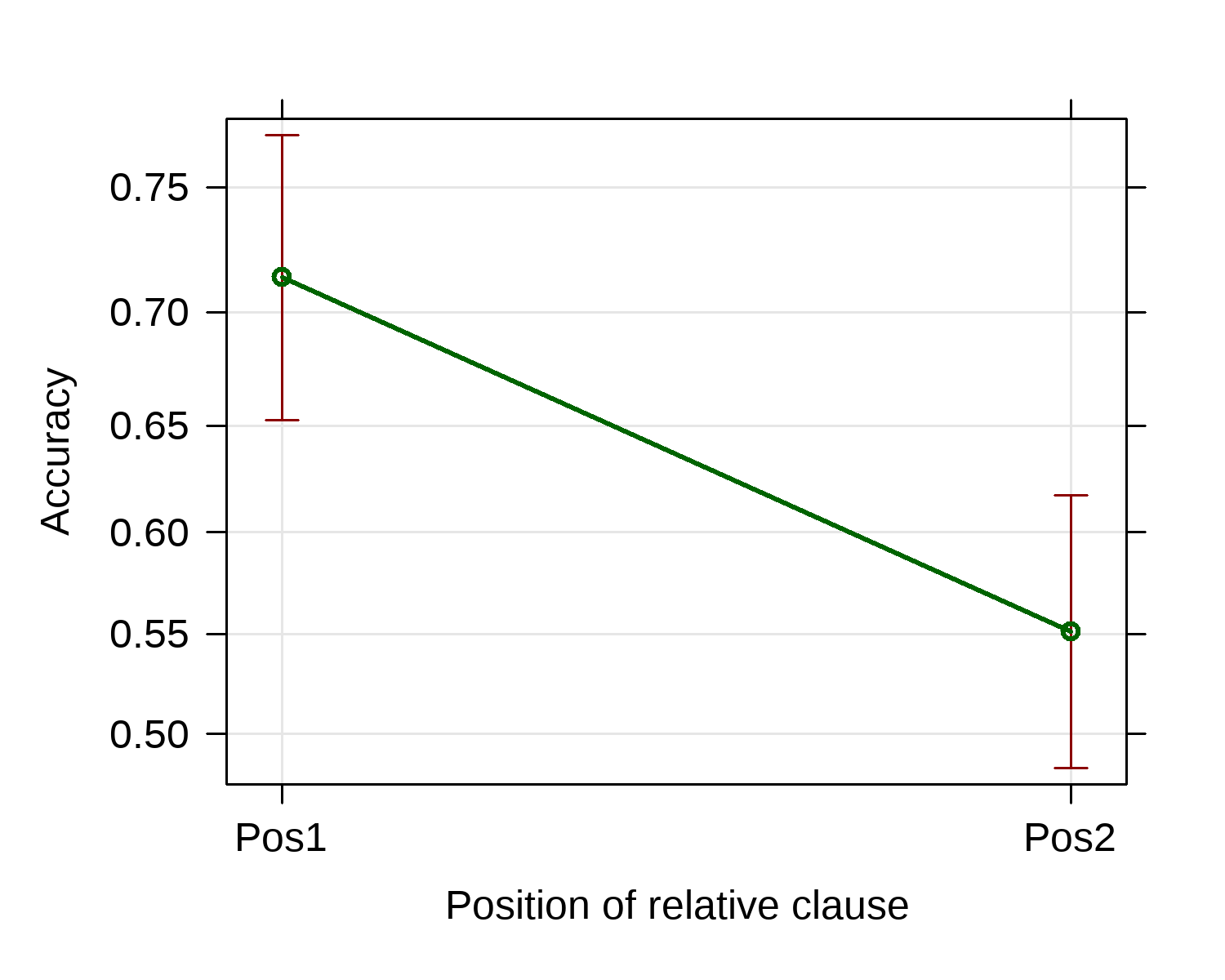


| Hauptsatz | Position | Anteil korrekter Antworten | | Reaktionszeit bei Frage | |
| --- | --- | --- | --- | --- | --- |
| beginnt mit |  | (Mittel) | (Stdev.) | (Mittel) | (Stdev.) |
| S | Pos1 | 0,74 | 0,44 | 3.713 | 2.114 |
| O | Pos1 | 0,63 | 0,48 | 3.826 | 2.202 |
| S | Pos2 | 0,62 | 0,49 | 4.002 | 2.354 |
| O | Pos2 | 0,46 | 0,50 | 4.142 | 2.404 |

The response times for the comprehension questions were longer for relative clauses in position 2 (mostly more than 4000 ms) than for those in position 1. The lengthening effect of constituent order on response time was more pronounced in object-initial main clauses (values above 4000 ms or just below) than for the other conditions. Overall, however, the differences between response times do not seem to be systematically different as in the case of response accuracy.

### 5.2 Regression analysis

To test the hypotheses stated above, generalised linear mixed effects regression analyses were conducted in which the following predictors (fixed effects) were chosen for response accuracy (as dependent variable): question type (HS question vs. RS question), constituent position in the main clause (HS\_order: SO vs. OS), constituent position in the relative clause (RS\_order: SO vs. OS), constituent position in the sentence relevant to an HS question or an RS question (RE\_S vs. RE\_O), congruence of constituent position in the main and relative clause (congruent vs. incongruent) and the position of the relative clause in the main clause (position 1, position 2). The variables participant and item served as random effects. However, the latter was excluded from the regression models described below due to non-significance. For the regression analysis, 992 observations of 62 participants on 16 test units were available.

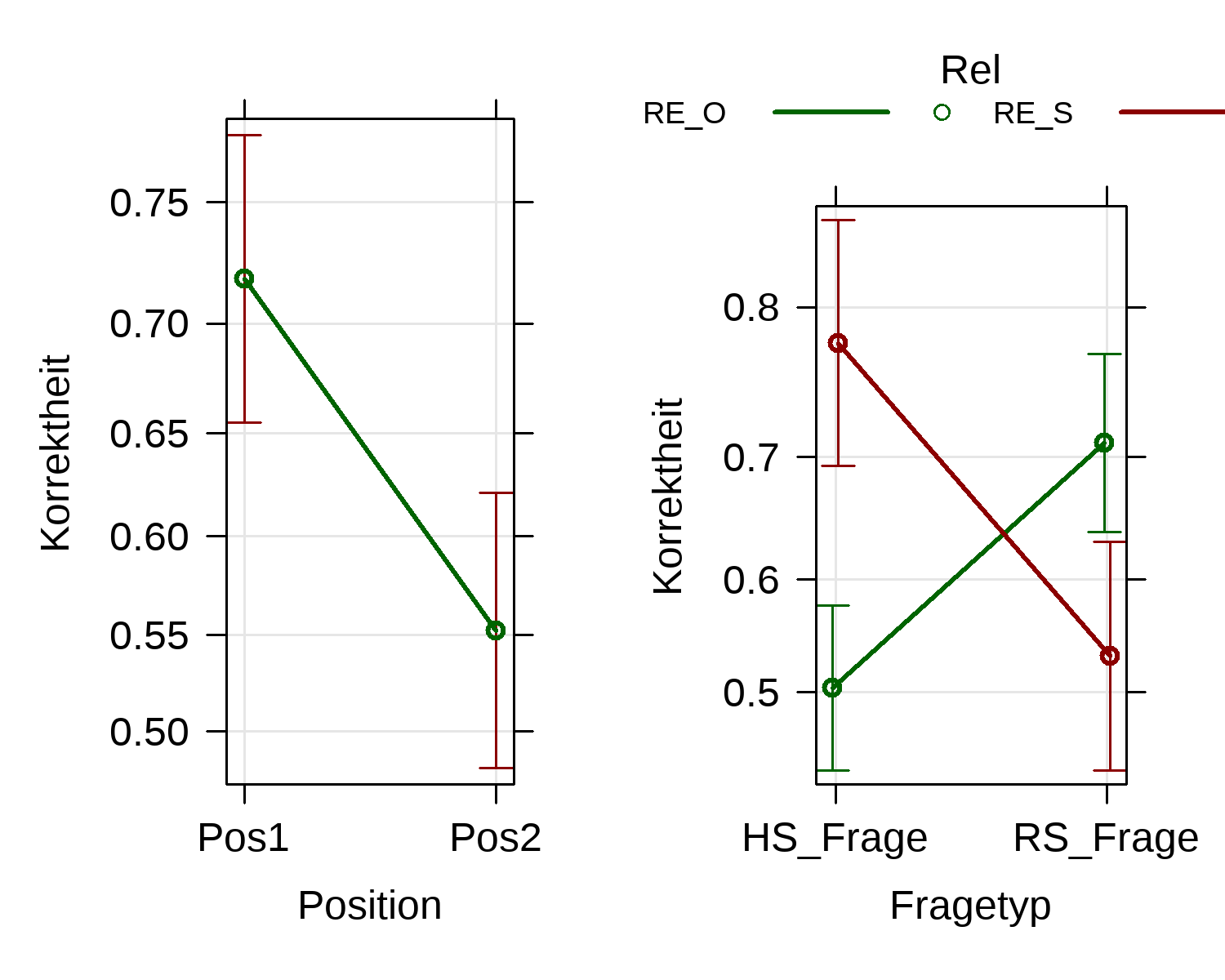


In the first regression model () two significant effects were identified, namely a main effect (position of the relative clause) and a three-way interaction between question type (HS question or RS question), constituent position in the main clause (subject- or object-initial) and congruence (i.e. whether or not the main and relative clauses both begin with a subject or an object). The difference between the intercept only null model (), in which none of the predictors were used, and model was statistically significant according to Akaike’s information criterion (AIC) ([Crawley](#ref-crawley2012r): 415): AIC = 1228.1; compared to the higher and thus less favourable value of the null model (AIC = 1285.3). The variance of response accuracy explained by model was = 32% (about 9% of the variance was explained by the assumed predictors, but the larger part, 23%, was explained by person-related differences). In model , in contrast to model below, no distinction was made between relevant and irrelevant clauses.

When a question was asked about the subject of the main clause (HS question), the proportion of correct answers was significantly greater (p < 0.001) for subject-initial main clauses (S\_initial) according to model than for objec-initial main clauses (O\_initial), both in the congruent and incongruent conditions. After an RS question (i.e., a question that focused on the subject of the relative clause), response accuracy was similar for both subject-initial and object-initial main clauses under the congruent condition, but less favourable than in the condition after an HS question. For incongruent clause pairs, response accuracy (as with HS questions) was significantly better for subject-initial main clauses than for object-initial ones. The response accuracy of the participants was therefore usually better when a main clause began with a subject. This finding is compatible with hypothesis . But it came as a surprise that congruent clause pairs actually reduce the accuracy of answers to RS questions.

The position of the relative clause after the first or second noun proved to be a significant main effect (p < 0.001), because the proportion of correct responses for relative clause in position 1 was greater than in position 2. This result confirms the hypothesis .

The interaction between the relative clause position and the constituent position in the main clause was not significant. This finding therefore does not confirm the hypothesis .



In the second regression model () a significant main effect (position) and an interaction effect (question type x constituent position in the relevant sentence; p < 0.001) were identified. In addition to a random intercept (differing from participant to participant), the model also included different slope coefficients for each participant. Model had a lower (better) AIC value (AIC = 1215.8) than model (AIC = 1226.9). Model explains = 37% of the variance of response accuracy, ie. more than model . This is mainly because model explains more intersubject variance (about 28%) and not because of the bigger contribution of the predictors in model .

In model , a distinction is made between relevant and irrelevant sentences for answering the respective comprehension question. If a question was asked about the subject of the main clause, then the answer accuracy for subject-initial main clauses relevant to the question was significantly higher (p = 0.001) than for object-initial relevant main clauses. If, on the other hand, a question was asked about the subject of the relative clause, then the answer accuracy was (surprisingly) better for object-initial relevant relative clauses. The proportion of correct answers to the comprehension questions was significantly greater for relative clauses in position 1 than in position 2 (p < 0.001).

While subject-initial relevant main clauses were better understood when an HS question was raised (*Who hit?*), it was the other way round with RS questions (*Who saw?*). For example, the object-initial relative clause (12) was better understood than the subject-initial one (13). Both relative clauses were in the less favourable position 2. The difference in response accuracy would be predictable with both models: in model with the three-way interaction question type x initial (main clause) x (clause) congruence, in model with the two-way interaction question type x relevant clause.

1. Zweifellos hat der Mann *den Anwalt*, [ *den* der Clown gesehen hat], geschlagen. (SO\_OS, relative clause in position 2, correct = 69%, response accuracy above average)
2. Zweifellos hat der Mann *den Anwalt*, [ *der* den Clown gesehen hat], geschlagen. (SO\_SO, relative clause in position 2, correct = 44%, response accuracy under mean accuracy with RS questions)

In the next comparison with a relative clause in position 1, that test unit (15) was understood slightly worse than (14) with an RS question, but not significantly worse. Since the main and relative clauses in (15) are congruent, according to model a (non-significantly) lower response accuracy was to be expected compared to (14). According to model , the object-initial relative clause (14) is expected to perform better than the subject-initial relative clause (15) with an RS question.

1. Zweifellos hat *der Mann*, [ *den* der Clown gesehen hat], den Anwalt geschlagen. (S\_OS\_O, relative clause in position 1, correct = 74%, response accuracy above average)
2. Zweifellos hat *der Mann*, [ *der* den Clown gesehen hat], den Anwalt geschlagen. (S\_SO\_O, relative clause in position 1, correct = 68%, response accuracy above average)

For HS questions, the response accuracy was higher overall than for RS questions and especially better for subject-initial main clauses than for object-initial ones. For relative clauses in position 1, it did not matter much whether the relative clause was connected subject-initially or object-initially (cf. (16) and (17)). Response accuracy was poorest when an HS question was raised after a test unit in which the main clause or even both clauses were object-initial and the relative clause was in position 2 (cf. (18) and (19)). Both models predict this outcome.

1. Zweifellos hat *der Mann*, [ *der* den Clown gesehen hat], den Anwalt geschlagen. (S\_SO\_O, relative clause in position 1, correct = 79%, response accuracy above average)
2. Zweifellos hat *der Mann*, [ *den* der Clown gesehen hat], den Anwalt geschlagen. (S\_OS\_O, relative clause in position 1, correct = 76%, response accuracy above average)
3. Zweifellos hat den Mann *der Anwalt*, [ *der* den Clown gesehen hat], geschlagen. (SO\_SO, relative clause in position 2, correct = 42%, response accuracy below average with HS questions)
4. Zweifellos hat den Mann *der Anwalt*, [ *den* der Clown gesehen hat], geschlagen. (SO\_OS, relative clause in position 2, correct = 37%, response accuracy below average with HS questions)

The regression model confirms the hypotheses and , but not (as does model ). Overall, subject-initial sentences were better understood by our participants than object-initial ones, and relative clauses after the first noun of the main clause (position 1) were better understood than those after the second noun of the main clause (position 2). However, model also shows a limitation of hypothesis : though it could be fully confirmed in main clauses, it could not in relative clauses. When answering questions about the subject of the relative clause (RS questions), object-initial relative clauses even seemed to be more favourable for our participants. However, an interaction between constituent position in the main clause and relative clause position, which was predicted according to hypothesis , could not be confirmed.

## 6 Conclusion

In an experiment to measure the processing effort of sentence comprehension, preference hypotheses on word order and relative clause position were tested in German as a second or foreign language. The experiment involved 62 Slovenian students of German studies in Maribor as participants. The linguistic material consisted of 16 utterances, each with a main clause and a relative clause. In each of the two sentences, either the subject could come before the object or vice versa. The relative clause could occur either after the first nominal phrase or after the second. The procedure used was the self-paced reading paradigm. Each word appeared individually on the screen, and the participant determined when the next word appeared on the screen by pressing a button. After each test unit, the participants were asked a comprehension question, either about the subject of the main clause (*who hit?*) or about the subject of the relative clause (*who saw?*). The test units appeared in randomised order.

The experiment tested the following hypotheses:  
() Comprehension questions about a subject-initial clause (SO-sequence) are answered correctly more often than those about an object-initial clause (OS-sequence). This hypothesis is based on the typologically founded preference of subject-initial sentences.

() Comprehension questions about a sentence including a relative clause in the first position (i.e. after the first noun in the main clause) are answered correctly more often than those with a relative clause in the second position (i.e. after the second noun in the main clause). This hypothesis is based on the assumption that a greater burden is placed on the working memory when processing a late appearing relative clause, since one has to remember several potentially relevant referents.

() Comprehension questions on a subject-initial sentence with a relative clause in position 1 are answered correctly more often than those on subject- or object-initial sentences with a relative clause in position 2. The interaction between syntactic relation and position of the relative clause is expected due to the working memory load, the preference for subject-initial sentences described above, and the implicative relativisability scale for syntactic relations.

The hypotheses were tested in a generalised mixed-effects regression analysis. The best regression model confirmed the hypotheses and , but not . Overall, subject-initial sentences were better understood by our participants than object-initial ones, and relative clauses after the first noun of the main clause were better understood than those after the second noun of the main clause. However, the hypothesis could be fully confirmed in main clauses (regardless of question type), but not in relative clauses when participants being asked about the subject of the relative clause. Furthermore, an interaction between constituent position in the main clause and relative clause position, which was predicted according to hypothesis , could not be confirmed.

## References

Abraham, Werner. *Deutsche Syntax Im Sprachenvergleich*. Narr, 1995.

Bartoń, Kamil. *MuMIn: Multi-Model Inference*. 2020, <https://CRAN.R-project.org/package=MuMIn>.

Bates, Douglas, et al. “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software*, vol. 67, no. 1, 2015, pp. 1–48, doi:[10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).

Bošković, Željko. “On Second Position Clitics Crosslinguistically.” *Formal Studies in Slovenian Syntax: In Honor of Janez Orešnik*, 2016, pp. 23–53.

Comrie, Bernard, and Tania Kuteva. “Relativization on Subjects.” *The World Atlas of Language Structures Online*, edited by Matthew S. Dryer and Martin Haspelmath, Max Planck Institute for Evolutionary Anthropology, 2013, <https://wals.info/chapter/122>.

Crawley, Michael J. *The r Book*. John Wiley & Sons, 2012.

Dressler, Wolfgang U., et al. *Leitmotifs in Natural Morphology*. Benjamins Amsterdam, 1987.

Dryer, Matthew S. “Order of Subject, Object and Verb.” *The World Atlas of Language Structures Online*, edited by Matthew S. Dryer and Martin Haspelmath, Max Planck Institute for Evolutionary Anthropology, 2013, <https://wals.info/chapter/81>.

Forster, Kenneth I., and Jonathan C. Forster. “DMDX: A Windows Display Program with Millisecond Accuracy.” *Behavior Research Methods, Instruments, & Computers*, vol. 35, no. 1, Springer, 2003, pp. 116–24.

Fox, Barbara A. “The Noun Phrase Accessibility Hierarchy Reinterpreted: Subject Primacy or the Absolutive Hypothesis?” *Language*, JSTOR, 1987, pp. 856–70.

Franks, Steven, and Tracy Holloway King. *A Handbook of Slavic Clitics*. Oxford University Press, 2000.

Gibson, Edward. “Linguistic Complexity: Locality of Syntactic Dependencies.” *Cognition*, vol. 68, no. 1, Elsevier, 1998, pp. 1–76.

Gohel, David. *Flextable: Functions for Tabular Reporting*. 2021, <https://CRAN.R-project.org/package=flextable>.

Golden, Marija, and Milena Milojević Sheppard. “Slovene Pronominal Clitics.” *Clitic Phenomena in European Languages*, John Benjamins Amsterdam, 2000, pp. 191–208.

Greenberg, M. *Short Slovene Reference Grammar. Reference Grammar Network. Duke University/University of North Carolina, Chapel Hill: SEELRC*. 2006, <http://www.seelrc.org:8080/grammar/pdf/stand_alone_slovene.pdf>.

Haftka, Brigitta. “Deutsch Ist Eine v/2-Sprache Mit Verbendstellung Und Freier Wortfolge.” *Deutsch-Typologisch*, de Gruyter, 2019, pp. 121–41, <https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/8937/file/Hafka_Deutsch_ist_eine_V_2_Sprache_1996.pdf>.

Haider, Hubert, and Luka Szucsich. “Slavic Languages–" SVO" Languages Without SVO Qualities.” *Ms. Submitted. Downloadable*, 2018.

Junker, Martina. *Der Verflixte Akkusativ: Altersunterschiede Und Altersinvarianz Beim Verstehen von sätzen Mit Unterschiedlich Komplexer Syntaktischer Struktur*. Universität Potsdam, 2004.

Just, Marcel A., et al. “Paradigms and Processes in Reading Comprehension.” *Journal of Experimental Psychology: General*, vol. 111, no. 2, American Psychological Association, 1982, p. 228.

Kaan, Edith. *Processing Subject-Object Ambiguities in Dutch*. Rijksuniversiteit Groningen, 1997.

Kaiser, Elsi, and John Trueswell. “Putting Things in Context: Sentence Processing in Languages with Flexible Word Order.” *University of Pennsylvania Working Papers in Linguistics*, vol. 9, no. 1, 2003, p. 9.

Keenan, Edward L. “Variation in Universal Grammar.” *Analyzing Variation in Language: Papers from the Second Colloquium on New Ways of Analyzing Variation*, vol. 2, Georgetown University Press, 1975, p. 136.

Keenan, Edward L., and Bernard Comrie. “Noun Phrase Accessibility and Universal Grammar.” *Linguistic Inquiry*, vol. 8, no. 1, JSTOR, 1977, pp. 63–99.

Marušič, Franc. “Positioning Slovenian Clitics.” *Ms. University of Nova Gorica*, Citeseer, 2009.

Nerantzini, Michaela, et al. “Processing Voice Morphology and Argument Structure by Greek Beginning Readers and Children with Reading Difficulties.” *Applied Psycholinguistics*, vol. 43, no. 1, Cambridge University Press, 2022, pp. 125–55.

Ozeki, Hiromi, and Yasuhiro Shirai. “Does the Noun Phrase Accessibility Hierarchy Predict the Difficulty Order in the Acquisition of Japanese Relative Clauses?” *Studies in Second Language Acquisition*, vol. 29, no. 2, Cambridge University Press, 2007, pp. 169–96.

Petrič, Teodor. “Zur Aneignung Der Verbstellungen Bei Slowenischen Deutschlernern.” *Vestnik*, vol. 32, no. 1-2, Društvo za tuje jezike in književnosti Slovenije, 1998, pp. 255–68.

Shenoy, Avinash R. *Grafify: A Package for Easy Grafs, ANOVAs and Posthoc Comparisons*. 2021, <https://github.com/ashenoy-cmbi/grafify>.

Tomlin, Russell S. *Basic Word Order (RLE Linguistics b: Grammar): Functional Principles*. Routledge, 2014.

Toporišič, Jože. *Slovenska Slovnica*. Obzorja, 1992.

Weyerts, Helga, et al. “Word Order in Sentence Processing: An Experimental Study of Verb Placement in German.” *Journal of Psycholinguistic Research*, vol. 31, no. 3, Springer, 2002, pp. 211–68.

Wickham, Hadley, et al. “Welcome to the tidyverse.” *Journal of Open Source Software*, vol. 4, no. 43, 2019, p. 1686, doi:[10.21105/joss.01686](https://doi.org/10.21105/joss.01686).