Ambiguity in case marking does not affect the description of transitive events in German: evidence from sentence production and eye-tracking

Judith Schlenter a, Yulia Esaulova a, Sarah Dolscheid b and Martina Penke a

aDepartment of Special Education and Rehabilitation, Faculty of Human Sciences, University of Cologne, Cologne, Germany; bBiopsychology and Cognitive Neuroscience, Faculty of Psychology and Sports Science, Bielefeld University, Bielefeld, Germany

ABSTRACT
The current study examined how German speakers described a scene where an agent acts upon a patient when the patient of the event was cued (a red dot preceding the patient, Experiment 1 vs. preview of the patient, Experiment 2). Prior research has shown that effects of attention manipulation on syntactic choice display cross-linguistic variation with notable differences between languages that have morphological case marking on noun phrases and English that lacks such marking. Since in German nominative subject case and accusative object case are unambiguously marked on masculine nouns but not on feminine nouns, it provides the ideal testing ground to investigate how case marking affects sentence production. Our results did not reveal any effect of case marking although the different types of attention manipulation were effective. Moreover, the eye-gaze data revealed that German speakers applied the same sentence-planning strategy for both masculine nouns (unambiguous) and feminine nouns (ambiguous).

ARTICLE HISTORY
Received 22 December 2020
Accepted 28 December 2021

KEYWORDS
Sentence production; visual cueing; eye-tracking; morphological case marking; non-canonical sentences; passive voice

Introduction
When describing an event where an agent acts upon a patient, speakers of English have the possibility to alternate between an active voice utterance (e.g. “The pirate pushes the parrot”) and a passive voice utterance (“The parrot is pushed by the pirate”). However, the first option is the one predominantly chosen, particularly when the agent of the action is animate, and the patient is inanimate or non-human. One reason for this is that animate entities are conceptually more accessible than inanimate entities (e.g. Bock et al., 1992; Esaulova et al., 2020; Prat-Sala & Branigan, 2000). Animacy is one of several prominence-lending features that can affect both a speaker’s attentional state and syntactic choice (Myachykov et al., 2018b; see also Bornkessel-Schlesewsky & Schlesewsky, 2009). Previous work has revealed close links between the allocation of attention and sentence production (e.g. Gleitman et al., 2007; Myachykov, 2007; Myachykov et al., 2012; Myachykov et al., 2018a). For instance, when English speakers’ attention was directed towards the patient referent in an event scene, speakers were more inclined to produce a passive sentence, hence starting their utterance with the referent in the spotlight of attention. These experiments showed that the presentation of a visual cue at the patient’s location prior to the onset of the event scene reduced the predominance of active voice and increased the number of passive voice utterances. The method of visual cueing for attention orienting goes back to Posner (1980) and, since then, has been adapted in many psycholinguistic experiments, where it has proven successful in eliciting non-canonical structures such as passives; for a recent review see Myachykov et al. (2018b). Taken together, cueing one referent, more precisely its location, by means of a visual cue is an effective way of increasing the prominence of an upcoming referent and its accessibility. However, in comparison to English, speakers of other languages such as German differ in the extent to which this increase in referent accessibility affects speakers’ syntactic choice.

For German, Esaulova et al. (2019) found that visual cueing of the patient referent had no effect on syntactic choice when the cue was presented for 60 milliseconds (ms) and thus was subliminal/implicit (6% passives after patient-cueing). By contrast, an implicit cue yielded a substantial increase of passive voice utterances in English speakers (26% passives, Gleitman et al., 2007; Myachykov et al., 2011a). When the visual cue was presented for 700 ms and thus was explicit, German speakers’ production of passives increased to 13%. This increase in passives is in line with a study reported by
Myachykov et al. (2018a), where English speakers were about 8% more likely to produce a passive after explicit 700 ms patient-cueing than after implicit 70 ms patient-cueing. Yet, effects of attention manipulation on syntactic choice were stronger for English than German: explicit visual cueing of the patient resulted in a higher percentage of passives in English (around 25-30%, Myachykov et al., 2018a: Fig. 4) than in German (13%, Esaulova et al., 2021). In a nutshell, there is a quantitative difference as regards the effect of patient cueing on passive production between English and German and a qualitative difference, with the need for an explicit cue for German speakers to be affected by patient-cueing.

**The relationship between word order and morphological case marking**

What are the differences between English and German that may underlie these discrepancies? Although both are Germanic languages, English is often described as a language with fixed word order, while German allows for more flexibility in word order (e.g. Bader & Häussler, 2010; Bakker, 1998; Bornkessel et al., 2002; Primus, 1998; Suijter et al., 2021). For instance, changing the order of constituents in English results in a clear change in meaning (e.g. the man tickles the boy vs. the boy tickles the man). In contrast, a speaker of German could produce the abovementioned sentences with no change in meaning thanks to overt case marking ([(der Mann)NOM/Agent kitzelt [den Jungen]ACC/Patient vs. [den Jungen]ACC/Patient kitzelt [der Mann]NOM/Agent = “The man tickles the boy”). It is possible that differences in word order flexibility and case marking also underlie the observed differences between speakers of English and German during attentional cueing tasks. These two properties are closely intertwined: Typological studies show that languages with flexible word order often assign grammatical function by means of morphological case marking, whereas languages with fixed word order often lack case marking (e.g. Blake, 2001; Siewierska, 1998). A similar trade-off between case marking and word order flexibility has been attested during language learning: When two groups of learners were exposed to miniature artificial languages containing optional case marking and either fixed or flexible word order, learners of the version with flexible word order employed case marking significantly more often (Fedzechkina et al., 2017).

As these findings show, the use of morphological case marking tends to go hand in hand with a greater flexibility in word order. For that reason, it is hard to conclude what underlies the observed differences between English and German. Here, we took a new approach to shed light on this issue. Specifically, we conducted a within-language comparison in German to determine whether overt case marking plays a crucial role in visually cued sentence production. To achieve this goal, we took advantage of the fact that German marks case overtly for masculine noun phrases (NPs) but not for feminine NPs. That means that thematic roles are unambiguously marked for masculine nouns but not for feminine nouns. While previous studies on German exclusively focused on masculine nouns that require overt case marking (Esaulova et al., 2019, 2020; Esaulova et al., 2021), in the current study we included feminine nouns for which case marking is covert, thus allowing for a better comparison to English. We reasoned that if case marking was an important property contributing to the previously observed discrepancies between English and German speakers, we should see similar differences between German masculine NPs that require overt case marking and German feminine NPs that carry no overt case marking and hence more closely resemble full NPs in English.

**Effects of attentional cueing on syntactic choice in English versus case-marking languages**

Besides differences between English and German, research has also observed differences between English speakers and speakers of other case-marking languages such as Finnish (Myachykov, 2007; Myachykov & Garrod, 2008; Myachykov et al., 2011a), Russian (Myachykov, 2007; Myachykov & Garrod, 2008; Pokhoday et al., 2019) and Korean (Hwang & Kaiser, 2015: Experiment 2). Although speakers of these languages were found to differ in several ways when tested in attentional cueing tasks, they all have in common that they deviated from English in their propensity to produce passives when the patient was visually cued.

In order to explain these differences, a number of suggestions have been made. For instance, Tomlin and Myachykov (2015) proposed that languages may differ in how strongly grammatical role assignment or linear ordering is affected by visual cueing (see also Myachykov et al., 2011b). According to their “mapping mechanism account”, a speaker would first try to map the most accessible referent to the most prominent grammatical role, the subject. In cases where the patient is the most prominent entity, this would result in a passive (provided this option is available in a given language). If a voice-based alternative is not (or “less”) available because the resulting passive structure is highly dispreferred, it has been suggested for Russian (e.g. Myachykov et al., 2011b), linear ordering would be the second-best choice. Here
the visually prominent patient is realised in sentence-initial position, resulting in an active object-initial structure. In English that lacks overt case marking on NPs, active object-initial sentences are not an option. Therefore, grammatical role assignment that maps a visually prominent patient to the subject role, and linear ordering that places a visually prominent patient in sentence-initial position both lead to the same result: the production of a passive clause where the patient is realised as subject in sentence-initial position. Hence, the greater propensity to produce passives after patient-cueing in English may stem from the two mechanisms pushing in the same direction, potentially enhancing each other.1

An alternative account has been proposed by Hwang and Kaiser (2015). In their study, the authors compared sentence production after attentional cueing in speakers of English and Korean, a language with a more flexible word order and case marking. Hwang and Kaiser found that Korean speakers’ syntactic choice was not modulated by implicit visual cueing contrary to English speakers. Overall, Korean speakers produced more passives than object-initial sentences, in fact even more than English speakers, although the Korean passive is reported to be infrequent and dispreferred.2 To explain these findings, Hwang and Kaiser propose that the more structural alternatives are available to describe a transitive event, the less likely a speaker of a case-marking language may be influenced by referent accessibility (“syntactic flexibility account”). Hence, speakers of languages which mark grammatical function by morphological case show a reduced effect of visual cueing because of the availability of structural alternatives (see also Myachykov et al., 2013). The rigid relationship between word order and grammatical function in English, on the other hand, makes it easier to assign a grammatical function to the most perceptually prominent and thus more accessible referent at an early stage of sentence planning. Supporting this, Hwang and Kaiser (2014) found that Korean speakers, unlike English speakers (Ferreira, 1996), showed a slow-down in sentence production when they could choose between alternative grammatical function assignments.

In addition to the outlined accounts, differences in sentence planning could explain why languages with and without morphological case marking are differently affected by attentional cueing. A common assumption in the production literature is that the planning process is divided into at least two processes (Levelt, 1989): a conceptualisation process, involving the generation of the preverbal message, and a formulation process. During the formulation process, a speaker accesses and selects lemmas and specifies the grammatical relations to finally convey who did what to whom. However, it is debated how far speakers plan ahead, that is, what is the size of the chunks or increments delivered to the linguistic encoding stage. Against this backdrop, two different sentence-planning strategies have been discussed: the hierarchical incremental strategy (e.g. Griffin & Bock, 2000; Bock et al., 2004) and the linear incremental strategy (Gleitman et al., 2007). Following a linear incremental strategy, a speaker starts with the encoding of the most accessible concept, which then becomes the first increment in the sentence to be uttered. The first lemma retrieved from the mental lexicon constrains the structure of the sentence, and sentence formulation proceeds incrementally as further lemmas are retrieved. The linear incremental strategy can also be described as word-by-word or phrase-by-phrase planning. Following a hierarchical incremental strategy, a speaker starts with the relational encoding of the event. Thus, planning entails more than just the first increment, namely some kind of information about the sentence structure (Bock et al., 2004), meaning that a structural representation guides lemma retrieval.

It has been suggested that eye-tracking analyses may indicate whether speakers follow one strategy or the other. During linear planning, speakers would rapidly fixate the character to be mentioned first and keep fixating the first mentioned character until approximately speech onset (Gleitman et al., 2007). Konopka et al. (2018, p. 74) see evidence for hierarchical planning in “a clearer temporal separation between message-level and sentence-level encoding during early formulation”. On the basis of prior research (Griffin & Bock, 2000), they assume that within the initial 400 ms after scene onset speakers distribute their gaze between the characters in the scene, indicating a phase of relational encoding of the event, whereas after around 400 ms they start to fixate the characters in their order of mention. Researchers have suggested that the necessity of early morphological marking on sentence constituents might invoke hierarchical planning (e.g. Norcliffe & Konopka, 2015, pp. 84–87). If the grammatical function of a noun has to be indicated by case marking, this may require advance planning of the structural relation between characters. However, this assumption remains speculative as, so far, only an influential role of verb morphology has been attested. For example, Norcliffe et al. (2015) found that speakers of Tzeltal, a language with morphologically complex verbs, made use of hierarchical planning when producing verb-initial sentences and linear planning when producing subject-initial sentences (for findings from the verb-initial language Tagalog, see Sauppe et al., 2013). The data suggest that speakers can flexibly apply one of both strategies depending on the necessity of morphological marking.
After patient-cueing, a speaker of English, a language which assigns grammatical functions by word order, may be more likely to apply a linear incremental planning strategy and to focus immediately on the patient to be realised in sentence-initial position (Gleitman et al., 2007). In contrast, a speaker of a case-marking language may be more likely to focus on the relation between agent and patient, employing a hierarchical incremental strategy, because the necessity to mark case requires the identification of thematic roles and linking these to syntactic functions. Thus, speakers of case-marking languages might not show an immediate preference for one referent over the other in early scene inspection.

The present study

The primary interest of the present study was to investigate how morphological case marking on NPs affects the description of transitive events during attentional cueing. For this, we capitalised on case marking and case syncretism in German, and systematically varied the gender of agent and patient. In German, nominative subject case and accusative object case are distinctly marked on the determiner for masculine nouns (der[NOM] vs. den[ACC]), as illustrated in example (1-a). In contrast, due to case syncretism in German, for feminine nouns nominative and accusative are not distinctly marked. In example (1-b) the same determiner die precedes the subject and object. Hence, sentences with two feminine nouns are syntactically ambiguous between active SVO or OVS.

If morphological case marking contributed to the previously observed discrepancies between English and German speakers, we would expect that visual patient cueing leads to more non-canonical sentences for feminine nouns than for masculine nouns. Critically, feminine NPs in German more closely resemble full NPs in English. Consequently, feminine nouns in German. Crucially, feminine NPs in no di-rect case marking played only a secondary role, then any difference observed for feminine nouns would be expected to be smaller than for masculine nouns. Alternatively, if cueing leads to more non-canonical sentences for feminine nouns, we would expect that visual patient cueing would lead to more non-canonical sentences for feminine nouns than for masculine nouns. This would be consistent with the idea that feminine NPs in German may be more likely to be realised in sentence-initial position (Gleitman et al., 2007).

The study we report in this paper, for the first time, compared the production of non-canonical structures in German and English. The two experiments differed in the type of attention manipulation. In Experiment 1, a simple cue, a red dot, was shown at the position where the patient was to appear in the event scene. In Experiment 2, speakers were given a preview of the patient, henceforth referred to as referential cueing. Hence, speakers could already activate the concept of the patient – and its case-related gender information – in their mental lexicon before the event scene appeared (for discussion, see Esaulova et al., 2020; Myachykov et al., 2012; Myachykov et al., 2018a). While we assumed that both cues increase the prominence of the patient and thus its accessibility, we assumed that both tap into different planning stages: a stage before any lemma selection and lexical access has taken place (simple cue) and a later stage where at least the concept of the patient has been activated (referential cue).

Experiment 1

In Experiment 1, we cued the patient of the upcoming event scene by means of an explicit visual cue, a red dot shown for 700 ms. Moreover, we always presented an animate patient to the left of an animate agent as
previous research had shown that in addition to animacy, the visuo-spatial position of patients relative to agents increases the prominence of the patient and thus the production of non-canonical structures in German (Esaulova et al., 2019; for a replication in Russian, see Pokhoday et al., 2019). If ambiguity in case marking were to affect the production of non-canonical sentence structures, we would expect to elicit fewer non-canonical structures for masculine nouns than for feminine nouns. Alternatively, if ambiguity in case marking were irrelevant to the production of non-canonical sentence structures, there should be no difference between masculine nouns and feminine nouns. Recall that for masculine nouns the speaker’s syntactic choice is set as soon as the sentence-initial NP is uttered (or the speaker has to restart) as morphological case on the sentence-initial NP has to be marked distinctively.

In line with our expectations for syntactic choice, we expected speakers to be faster in producing sentences with feminine nouns if they postponed grammatical role assignment and proceeded with the production of the first sentential increment. Alternatively, if ambiguity in case marking were not affecting sentence planning, there should be no difference in sentence-planning latencies between masculine nouns and feminine nouns. A difference in sentence planning between masculine nouns and feminine nouns should also be reflected in speakers’ eye-gaze behaviour. As outlined previously, morphological case marking on the sentence-initial NP might lead to hierarchical incremental planning. Thus, during the first 400 ms after scene onset, fixations on agent and patient might be equally distributed and a preference for the first-mentioned character only emerges towards speech onset. Since in German case is unambiguously marked for masculine nouns but not for feminine nouns, German speakers might apply a hierarchical incremental planning strategy for masculine nouns but not for feminine nouns. For feminine nouns, German speakers might show an immediate preference for the first-mentioned character in line with linear incremental planning. Alternatively, German speakers might apply the same planning strategy, either a hierarchical or a linear incremental strategy, for both masculine nouns and feminine nouns.

**Methods**

**Participants**

Forty-three native speakers of German recruited from the student population at the University of Cologne took part in this study. Two participants were excluded due to difficulties with calibration for eye-movement recording. In total, 41 participants (31 female, mean age: 22.83 years, SD = 3.41 years) were included into the statistical analysis. Participants had normal or corrected-to-normal vision and reported no language or attention-related impairments. They received a course credit or monetary remuneration for their participation. Written informed consent in accordance with the Declaration of Helsinki was obtained from all participants at the beginning of the experimental session. The protocol was approved by the Ethics Commission of the Faculty of Medicine at the University of Cologne (approval number 16-134).

**Materials and Design**

A set of 24 black-and-white drawings that depicted a transitive event were created for the experiment. Each event scene showed a human agent acting upon a human patient. The patient was always displayed to the left of the agent. In half of the event scenes both the agent and the patient were female (see Figure 1, A), and in the other half both of them were male (see Figure 1, B). Twelve feminine and 12 masculine nouns matching in the number of syllables ($M = 2$, range: 1-3) were selected as agents and patients. They did not differ in frequency ($M_{feminine} = 28.92$, $M_{masculine} = 37$, $t(18.42) = 0.41$, $p = 0.684$). As regards morphological complexity, no compounds were used, however, ten nouns (five feminine and five masculine) included productive derivations (e.g. *König*-*in* “queen”). Each noun appeared once as an agent and once as a patient. The depicted characters were similar in size, visual complexity, and distance, and there was no contact between them. Six verbs were selected based on visual recognisability as a prototypical transitive event (filmen “film”, schieben “push”, schlagen “hit”, geben “water”, messen “measure”, wiegen “weigh”). Each verb was presented four times in combination with different nouns. The verbs were comparable with regard to their likelihood to appear in active and passive voice frames (see Esaulova et al., 2019). All experimental items can be viewed at [https://osf.io/jwgk4/](https://osf.io/jwgk4/).

Twenty-four drawings depicting animals and inanimate objects that were situated next to each other were used as filler items (12 nouns with feminine gender, 12 nouns with masculine gender). The two figures were not involved in any action, so they could be described in locative sentences such as *The apple is next to the fish* (see Figure 1, C). The fillers were included to prevent participants from developing a strategy to repeat the same syntactic structure from trial to trial. Twenty-four picture familiarisation trials (in total) served to ensure that participants could easily recognise
the characters and objects that would later appear in the test trials. They consisted of 24 displays showing four pictures each, a female and a male character and an animal and an inanimate object (see Figure 1, D).

The experimental design included the factor case (unambiguous/masculine vs. ambiguous/feminine) as within-subjects and between-items factor. There was a single presentation list split into six blocks that consisted of four experimental and four filler items. Each block was preceded by four familiarisation trials. The order of test trials within blocks was randomised.

Procedure

Participants were seated within a distance of 60 cm in front of an LCD monitor connected with an EyeLink 1000 Plus eye tracker. Viewing was binocular but only the dominant eye was tracked (sampling rate 500 Hz). To ensure the accuracy of eye-movement recordings, a calibration and a validation were performed when the experiment started and repeated whenever necessary. During the experiment, participants were equipped with a wired headset with a boom microphone that served to provide auditory instructions and allowed speech recordings with a sampling rate of 50–5000 Hz.

Each experimental session started with a pre-recorded instruction. During the instruction, participants were given two visual examples, one for an experimental and one for a filler item, together with possible one-sentence descriptions (active voice, passive voice and object topicalization for the event scene and a locative sentence for animal and inanimate object). Participants were also instructed to assign either der Mann “the man” or die Frau “the woman” to characters they could not remember the names for. The instruction was followed by calibration/validation and a practice block that contained two practice items to ensure that participants understood the task of the experiment. After completing the practice block, the participants proceeded with the actual experiment. During the familiarisation, participants saw a screen displaying four characters/objects (e.g. bride, pirate, fish, apple) and were asked to indicate where the character/object is located by pressing the corresponding arrow key (top, right, bottom, left) on the keyboard. For example, after hearing “Where is the pirate?”, they had to indicate the position of the pirate. If the response was incorrect, the question was repeated, if it was correct, they were asked for another character/object in the visual display. This ensured that participants recognised the depictions of characters and objects as intended.

In the test trials, first a fixation cross (500 ms) appeared in the centre of the screen, followed by a red dot subtending 1° of visual angle for a duration of 700 ms. For the experimental items (24 in total) the red dot appeared in the centre of the left half of the screen, for the filler items (24 in total) the red dot appeared in the right half of the screen. Afterwards an experimental or a filler item was presented for 6000 ms; see Figure 2. Participants were instructed to describe the depicted item in one sentence.

A drift correction screen appeared after 1000 ms of a blank screen, which indicated the end of the trial and allowed participants to complete their utterance. When the participant’s fixation landed on a target point in the centre of the screen, the experimenter continued to the next trial by button press or restarted the calibration/validation procedure. An experimental session lasted approximately 30 minutes.

Data analyses

The data were analysed with respect to three measures: syntactic choice, that is, which utterance type participants produced, sentence-planning latency as measured by speech onset times (SOTs), and eye gaze. The total of 984 produced utterances for the experimental items were coded as active SVO, active OVS, or passive. Sentences that described no transitive event, sentences where participants used another gender for a noun than intended, incomplete sentences, trials with sentence-initial self-corrections or a distraction (e.g. talking to the experimenter) were excluded; for an overview of excluded trials, see Table S1 in the Supplementary...
Materials. This comprised 48 trials (4.9% of the data), leaving 936 trials for the analyses. To analyse sentence-planning latencies for the trials left after exclusion, we used a Praat (Boersma, 2001) script to determine the onset of each utterance. Where necessary, onsets were corrected by hand (e.g. in case of natural noise before speech production like uhm etc.). Since the audio recording of each utterance started at the beginning of each trial, we subtracted the time for the fixation cross (500 ms) and the time the perceptual cue was shown (700 ms) to obtain the time participants needed for sentence planning from the point in time the event scene was displayed on the screen. Since raw reaction times are rarely normally distributed, the Box–Cox procedure (Box & Cox, 1964; Osborne, 2010) was applied to determine the optimal transformation for the dependent variable SOT. For the eye-tracking analyses, two equally sized interest areas for the patient and agent were defined. The focus of our analyses was on the initial 400 ms time window as this time window has been reported critical for message generation in previous eye-tracking research (e.g. Griffin & Bock, 2000; Konopka et al., 2018; van de Velde et al., 2014). In line with a procedure typically applied for eye-tracking-during-listening experiments, we shifted the onset of our analysis window 200 ms forwards to account for the time it takes to initiate a saccade (Matin et al., 1993). Saccadic eye-movements as part of a gaze shift are associated with a shift of attention (e.g. Tanenhaus & Trueswell, 2006). Note that if speakers were fixating the patient at scene onset, as we expected them to do as a result of the visual cueing, they should maintain fixating the patient or, if their gaze was somewhere else, shift their focus of attention towards the patient when following a linear incremental strategy for the production of a patient-initial sentence. Likewise, they should shift their focus of attention towards the agent (or maintain fixating the agent) when following a linear incremental strategy for the production of an agent-initial sentence. To check whether effects of cueing were still visible at scene onset, we determined the proportion of first fixations on patients. Following a reviewer’s suggestion, we also provide information about early fixation sequences in the Supplementary Materials.

For statistical analyses, we computed mixed-effects models in R (R Core Team, 2020), using the lme4 package (Bates et al., 2015). To obtain p-values, we used the lmerTest package (Kuznetsova et al., 2017). The best-fitting model was selected based on the lowest AIC value of the converging models with random effect structure (Matuschek et al., 2017) varying from minimal to maximal (Barr et al., 2013). The data files and the R code for all models reported in the paper are available on OSF at https://osf.io/5wkc8/.

Results

Syntactic choice

For Experiment 1, participants almost exclusively produced SVO active sentences. Notably, there were no active voice utterances with an OVS structure. Table 1 shows the total number of utterance types and percentages.

We analysed whether non-canonical sentence production, a binary coded variable, was modulated by case marking as the independent categorical variable. Note that we had hypothesised that if ambiguity in
case marking affected non-canonical sentence production, speakers should produce more non-canonical structures for feminine nouns. The results from a mixed-effects logistic regression model, see Table 2, shows that the gender and thus case marking had no effect on syntactic choice: the production of passives did not differ between masculine nouns and feminine nouns. The factor case was treatment-coded with feminine as reference level. The significant effect at the intercept reflects that, for the respective reference level, active utterances outnumbered passive utterances.

**Sentence-planning latency**

Next, we analysed speech onset times to determine whether there was a difference in timing between active and passive voice utterances as well as a modulation by case marking. At least numerically, participants needed more time to produce a passive ($M = 1739.1$ ms, $SD = 416.85$) than an active utterance ($M = 1612.78$ ms, $SD = 474.86$), as further illustrated in Figure 3.

A linear mixed-effects model analysing the reciprocal square root transformed SOTs depending on voice choice (treatment-coded with active as reference level) and case marking (coded as 0.5/−0.5) confirmed that across feminine nouns and masculine nouns the difference between active and passive utterance onsets was significant ($\beta = -0.001$, $SE = 0.0004$, $t = -2.891$, $p = 0.006$). There was no effect of or interaction with case. Thus, irrespective of ambiguity in case marking speakers needed more time to produce a passive than an active utterance.

**Eye gaze**

Finally, we analysed speakers’ eye-gaze behaviour during sentence planning. The time course showing the proportion of fixations on patient and agent from scene onset is displayed in Figure 4. Due to the high number of actives (867 trials), the lines for active voice utterances at the top are smoother than for the passive voice utterances (69 trials).

For the active utterances, the majority of fixations at scene onset were on patients (61%), followed by other on-screen fixations (30%) and fixations on agents (9%). For the passive utterances, 59.5% of first fixations were on patients, followed by other on-screen fixations (39%) and fixations on agents (1.5%). Thus, after the presentation of the explicit visual cue at the patient’s location, in most cases participants’ visual attention at scene onset was still on the patient. Visually, the lines indicating looks on patient and agent for the active voice utterances start to diverge in another direction (away from the patient) after around 300 ms and remain on the agent until shortly before speech onset. For the passive voice utterances, speakers’ gaze rather remains on the patient with temporary shifts towards the agent after around 600 ms.

Mixed-effects logistic regression was used to analyse fixations of the patient and agent within each 20 ms time bin of every trial (unique participant-item-condition combination) for the time window between 200 and 600 ms after scene onset, which is indicated in Figure 4 by the two solid vertical lines. Note that we analysed looks on patient relative to looks on patient and agent, that is, no other on-screen looks were included. We start with the results from two separate analyses for active and passive utterances. The logistic regression model for the active voice utterances revealed that patient fixations were below chance level (i.e. 0 on the logit scale, corresponding to 0.50 on the probability scale). There was no effect of case, here coded as 0.5/−0.5 to get the grand mean of the two conditions for the intercept. The model output is summarised in Table 3. Thus, participants were less likely to fixate the patient than the agent of the event to be described at an early stage of sentence planning (before 600 ms) when formulating an active SVO sentence. A preference for the agent was found for feminine nouns (ambiguous case marking) and masculine nouns (unambiguous case marking) alike.

The logistic regression model for the passive utterances showed a preference for the patient over the agent, as indicated by the positive intercept term in Table 4. Like for the active utterances, there was no effect of case marking, that is, no difference between masculine nouns and feminine nouns.

Finally, we analysed fixations on patient and agent depending on the treatment-coded factor voice choice (note that voice choice was not a systematic manipulation). Since case had no effect in separate models for

---

**Table 1.** Contingency table showing the number of active and passive voice utterances for ambiguously case-marked feminine nouns and unambiguously case-marked masculine nouns in Experiment 1. Percentages are shown in parentheses.

| Case marking | Active (feminine nouns) | Passive (masculine nouns) | Total |
|--------------|-------------------------|---------------------------|-------|
| Active       | 448 (93%)               | 419 (92%)                 | 867 (93%) |
| Passive      | 32 (7%)                 | 37 (92%)                  | 69 (7%)  |

**Table 2.** Mixed-effects logistic regression model on the syntactic choice data in Experiment 1.

|                       | $\beta$ | SE   | z-value | p-value  |
|-----------------------|---------|------|---------|----------|
| Intercept (fem.)      | -8.919  | 1.614| -5.526  | < 0.001***|
| Case                  | 0.426   | 0.495| 0.862   | 0.389    |
Figure 3. Mean speech onset times and standard error bars for active and passive voice utterances shown for ambiguously case-marked feminine nouns (red solid line) and unambiguously case-marked masculine nouns (blue dashed line) in Experiment 1.

Figure 4. Proportion of fixations on patient (coloured) and agent (grey) for ambiguously case-marked feminine nouns (solid lines) and unambiguously case-marked masculine nouns (dashed lines) in Experiment 1. Since for the visualisation of the data we also included other on-screen looks, the proportion of fixations does not add up to one. The dotted vertical line indicates the average speech onset, the two solid vertical lines the analysis window.

Table 3. Mixed-effects logistic regression model on looks to the patient relative to looks on patient and agent for the active voice utterances in Experiment 1.

|                | \( \beta \) | SE   | z-value | p-value  |
|----------------|-------------|------|---------|----------|
| Intercept (fem./masc.) | −0.972 | 0.252 | −3.854 | <0.001*** |
| Case           | 0.164  | 0.286 | 0.572  | 0.567    |

Table 4. Mixed-effects logistic regression model on looks to the patient relative to looks on patient and agent for the passive voice utterances in Experiment 1.

|                | \( \beta \) | SE   | z-value | p-value  |
|----------------|-------------|------|---------|----------|
| Intercept (fem./masc.) | 3.294 | 1.084 | 3.038  | 0.002**  |
| Case           | 2.349  | 1.650 | 1.423  | 0.155    |
active voice and passive voice utterances, we did not include this factor as a covariate. The maximally specified model showed an effect of voice choice ($\beta = 3.848$, $SE = 0.986$, $z = 3.904$, $p = < 0.001$), with more patient fixations for passive utterances as compared to active utterances (intercept: $\beta = -0.957$, $SE = 0.249$, $z = -3.849$, $p = < 0.001$). Note though that the eye-gaze data for passive voice utterances, and hence also the comparison between utterance types, should be interpreted with caution due to the small number of passive observations entering the analyses.

Discussion

In the current experiment, we varied the gender of animate referents in event scenes (masculine/feminine) that determined their ambiguity in terms of case marking. Additionally, we cued the patient of the upcoming event scene by means of an explicit visual cue and presented patient characters to the left of animate agents in order to increase the production of non-canonical structures. We hypothesised that, if ambiguity in case marking affected syntactic choice, speakers should produce more non-canonical structures for feminine nouns as compared to masculine nouns. Similar to full NPs in English, NPs with feminine nouns in German lack overt case marking. Previous studies in English found an increase of passives after explicit patient cueing up to around 25-30% (Myachykov et al., 2012: Figure 4; Myachykov et al., 2018a: Figure 4). Moreover, we expected speakers to show faster sentence-planning latencies for feminine nouns as compared to masculine nouns and an immediate looking preference for the character to be mentioned first, following a linear incremental planning strategy rather than a hierarchical incremental strategy. The latter strategy we expected to find for masculine nouns that require overt and unambiguous case marking. Below we will discuss the results for syntactic choice, sentence-planning latency, and eye gaze in turn.

With respect to syntactic choice, the results showed no difference in the production of non-canonical structures between masculine nouns and feminine nouns and thus no effect of ambiguity in case marking. Overall, we elicited only a low number of passives and no object-initial active sentences, although both structures are available and felicitous in German. Thus, our manipulation of visual attention, here by an explicit visual cue presented prior to a left-positioned animate patient, did not increase the production of non-canonical sentences. Rather, the percentage of passive utterances produced in the current experiment, 7% overall, was similar to previous studies on German that used no visual cue (Sauppe, 2017b), an implicit cue (Esaulova et al., 2019) or an explicit 700 ms cue (Esaulova et al., 2021). The results showed that increasing the prominence of the patient in a scene-description task did not lead to similar increases in the number of passives produced as was observed for English. Moreover, the reluctance to deviate from canonical SVO structures was independent of the necessity to produce overt and distinct case markings.

In line with the findings for syntactic choice, analyses of SOTs showed no difference between masculine nouns and feminine nouns and thus no evidence for an effect of ambiguity in case marking on sentence-planning latency. Furthermore, the SOT results showed that across feminine and masculine nouns the planning of passive utterances needed more time than the planning of active utterances. The finding that sentence-planning latencies are longer for passives than actives supports previous findings for German native speakers and may indicate that the production of passives as compared to actives requires more cognitive resources (Esaulova et al., 2019; Sauppe, 2017b; see in particular Sauppe, 2017a), at least for animate patients (Esaulova et al., 2021).

The eye-gaze data indicated that, first, visual cueing was effective and in most cases attention at scene onset was drawn to the patient and, second, that characters were fixated in their order of mention (i.e. subject-object). We observed more looks to the agent than to the patient for active utterance trials and more looks to the patient than to the agent for passive utterance trials before 600 ms after scene onset. Consistent with the syntactic choice and the SOT results, speakers’ eye-gaze behaviour did not show any significant differences between masculine nouns and feminine nouns during the 200–600 ms time window. Critically, the eye-gaze data, most clearly the data for the active utterance trials, pointed to the application of a linear incremental strategy: Speakers rapidly fixated the character to be placed in sentence-initial subject position, a finding that is well in line with Gleitman et al. (2007). Only later, we noticed differences between active and passive utterances in the proportion of looks to agent and patient. Visual inspection of the time course graphs revealed that, after around 600 ms, speakers tended to shift their attention from patient to agent and back to the patient before they started to produce a passive utterance. This is consistent with the observation made by Sauppe (2017b), who noted that visual attention was more evenly distributed between characters during the planning of passive utterances. While Sauppe argued for weakly hierarchical planning in German based on this observation, we remain cautious about interpreting any eye-movement patterns emerging after 600 ms and
thus beyond the conceptualisation stage defined in prior research on sentence planning.

Altogether, Experiment 1 yielded four important findings: First, case marking did not affect sentence planning and syntactic choices. Second, if German speakers produced a non-canonical structure, they produced a passive sentence and no object-initial active sentence. Thus, German speakers in the current experiment, like in previous studies on German, made use of voice alternation to accommodate a prominent patient. Third, patient-cueing appeared to be less effective in German than in prior studies on English. Fourth, the eye-gaze behaviour pointed to linear incremental planning in German. Yet, Experiment 1 leaves a possibility that we see no evidence of case marking on scene descriptions for two reasons: A simple perceptual cue may not allow for an early enough activation of case information in a sentence-planning process (uninformative cue with respect to case-related referent features) and it elicited only a low number of non-canonical structures, which is problematic when analysing an effect of case marking on non-canonical sentence production. To rule out that the obtained results were influenced by these factors, we conducted Experiment 2.

**Experiment 2**

In Experiment 2, we changed the type of cue and used a referential cue. Recently, Esaulova et al. (2020) showed that priming participants with the picture of the patient referent further increased the production of passives in German (17.6% overall). Hence, their results indicated that perceptual prominence can be enhanced by conceptual priming, that is, activation of a concept, here the patient character. Participants in this study produced more passives in the patient preview condition (25%) than in the no-preview condition (11%). This effect of patient accessibility, however, was susceptible to character position and occurred for left-positioned but not for right-positioned patients. Moreover, SOTs were faster for passive utterances than active utterances, indicating that referential cueing facilitated the production of the non-canonical passive structure. Given these results, we conducted Experiment 2. Our reasons were twofold: A referential cue taps into a later stage of sentence planning than a simple perceptual cue we used in Experiment 1. We expected a referential cue to activate case-related information about the referent either on a conceptual level via the biological gender of the patient character (male or female) or on a grammatical level via concept-mediated activation of the lexical entry connected to a gender node. Moreover, we expected to elicit more non-canonical structures through referential cueing. Note that, unlike Esaulova et al. (2020), we presented the patient cue at the patients’ position, that is, in the left half of the screen, so our referential cue was also informative with regard to the patient’s visuo-spatial position.

**Methods**

**Participants**

Forty-five students from the University of Cologne took part in this study. Four participants had to be excluded, two due to eye-tracking problems, one because s/he turned out not to be naive with regard to the study aim and one because the majority of her/his produced utterances were sentences involving embedded non-finite constructions (x versucht y zu schlagen “x tries to hit y”). In total, 41 participants (35 female, mean age: 22.34 years, SD = 3.32 years) were included into the statistical analysis. All of them were native speakers of German who did not report any attention or language-related medical condition and had normal or corrected-to-normal vision. They received either a course credit or monetary remuneration for their participation. Written informed consent in accordance with the Declaration of Helsinki was obtained from all participants at the beginning of the experimental session. The protocol was approved by the Ethics Commission of the Faculty of Medicine at the University of Cologne (approval number 16-134).

**Materials and Design**

The materials and the design were the same as in Experiment 1.

**Procedure**

The procedure was analogous to Experiment 1 with the following, slight differences. After the fixation cross in the centre of the screen (500 ms), the (left-positioned) patient of the upcoming scene, or, in case of a filler item, the right-positioned animal/object was shown for a duration of 700 ms. The picture preview was followed by a 500 ms blank screen to avoid an animation effect. Finally, an experimental or filler item was presented (6000 ms) which had to be described in one sentence by the participants; for an illustration see Figure 5.

**Data analyses**

The same data analysis procedure as for Experiment 1 was applied. Forty trials (4.08% of the data) were excluded (for
more details, see Table S1 in the Supplementary Materials), leaving 944 trials for the analyses. For the SOT analyses, we subtracted from each onset previously determined in Praat the time for the fixation cross at the beginning of the trial (500 ms), the time the cue was shown (700 ms) and the time for the blank screen (500 ms), which was included in Experiment 2, to obtain the time participants needed for utterance production upon the presentation of an event scene.

**Results**

**Syntactic choice**

Like in Experiment 1, participants in Experiment 2 produced no OVS active sentences. However, we elicited considerably more passive utterances in Experiment 2 (32% as compared to 7% in Experiment 1). Table 5 shows the total number of utterance types and percentages for Experiment 2.

The results from a logistic regression model indicated that case marking did not affect passive production. There was no difference between feminine nouns (intercept: $\beta = -2.658$, $SE = 0.861$, $z = -3.088$, $p = 0.002$) and masculine nouns ($\beta = 0.147$, $SE = 0.524$, $z = 0.280$, $p = 0.78$). Since the referential cue in Experiment 2 visibly increased the production of passives, we further compared the data from both experiments to test for an effect of cue type.

For the joined dataset, we analysed whether case marking (ambiguous/feminine vs. unambiguous/masculine) interacted with cue type (simple vs. referential). To compare all levels with each other, we used the emmeans package (Lenth et al., 2020). Only comparisons between cue types showed significant differences (for all contrasts, see Table S.2 in the Supplementary Materials). Therefore, we only kept cue type as a fixed factor (treatment-coded) in the regression model. The (simple) effect of cue type in Table 6 indicates that participants indeed produced more passives after a referential cue than after a simple cue.

**Sentence-planning latency**

In Figure 6, we show the mean speech onset times for active voice and passive voice utterances per case condition in Experiment 2. Descriptively, SOTs in Experiment 2 were shorter for passives ($M = 1300.11$ ms, $SD = 599.02$) than actives ($M = 1542.88$, $SD = 547.8$).

A linear mixed-effects model analysing the fourth root transformed SOTs depending on voice choice (treatment-coded) and case (coded as 0.5/−0.5), however, showed that the difference between utterance choice...
types was not significant ($\beta_{\text{passive}} = -0.136, SE = 0.098, t = -1.378, p = 0.182$). There was no modulation by case.

When compared to Experiment 1, cue type appears to be a modulating factor. To test whether the effect of cue type was significant, we computed a linear mixed-effects model including both datasets, the results of which are given in Table 7. All factors were treatment-coded. While there was no effect of cue type for active utterances, the interaction between cue type and voice choice indicates that speakers in Experiment 2 needed less time to produce a passive, meaning that cue type affected sentence-planning latency for passives and led to facilitated production of passives after a referential cue.

### Eye gaze

The substantial increase of passives in Experiment 2 enabled us to better compare speakers’ eye-gaze behaviour during active and passive voice production. The proportion of fixations on agent and patient from scene onset is displayed in Figure 7.

There were 49% first fixations on patients for active utterances (645 trials) and 68% first fixations on patients for passive utterances (299 trials). First fixations on agents and other on-screen fixations were equally distributed for each utterance type (25.5% each for actives and 16% each for passives). We computed a logistic regression model analysing the counts of first fixations on patient versus elsewhere for Experiment 1 and Experiment 2 combined. While there was a trend towards more patient fixations for passive than for active utterances in Experiment 2, none of the other contrasts revealed a significant difference, that is, a difference between utterance types or between experiments (see Tables S.3-S.5 in the Supplementary Materials).

Next, we analysed looks on patient and agent for the 400 ms time window after scene onset, shifted 200 ms forwards to account for eye-movement latency. As for Experiment 1, we started with separate models for active voice and passive voice utterances. Table 8 shows the model output for actives and Table 9 the model output for passives. In both models, case was sum-contrasted using deviation coding (0.5/−0.5).

There was no effect of case neither for active utterances nor for passive utterances for the 400 ms time window. The intercept values indicate that, for both actives and passives, looks to patients were below chance level. In other words, speakers in Experiment 2 were more likely to fixate the agent than the patient. Whether patient fixations differed between active and passive voice utterances was tested in an additional model that only included voice choice as a factor (treatment-coded). A (simple) effect of voice choice showed that the difference between actives and passives was significant, with more looks to patients for passive utterances ($\beta = 1.925, SE = 0.832, z = 2.314, p = 0.021$) as compared to active utterances (intercept: $\beta = -3.747, SE = 0.483, z = -7.753, p = < 0.001$).

### Figure 6

Mean speech onset times and standard error bars for active and passive voice utterances shown for ambiguously case-marked feminine nouns (red solid line) and unambiguously case-marked masculine nouns (blue dashed line) in Experiment 2.

### Table 7

Mixed-effects linear regression model on speech onset times from Experiment 1 and Experiment 2.

|                      | $\beta$ | $SE$  | $t$-value | $p$-value |
|----------------------|---------|-------|-----------|-----------|
| Intercept (active, referential cue) | 6.211   | 0.059 | 104.815   | < 0.001*** |
| Cue Type (simple cue)      | 0.089   | 0.063 | 1.412     | 0.162     |
| Voice (passive)          | −0.137  | 0.087 | −1.586    | 0.124     |
| Cue Type:Voice           | 0.376   | 0.151 | 2.491     | 0.020*    |
To summarise the eye gaze results for the 400 ms time window for Experiment 2, similar to Experiment 1, there was no effect of case marking but a difference in looks to the patient between active voice and passive voice utterances. However, unlike in Experiment 1, speakers looked more at the agent than the patient during the 400 ms time window when producing a passive utterance.

In a next step, we analysed both experiments together. The overall eye-gaze pattern for active voice and passive voice utterances was similar in Experiment 1 and Experiment 2, as reflected by an effect of voice choice across cue types in a logistic regression model with sum contrast for cue type and active voice as reference level ($\beta_{\text{passive}} = 2.069$, $SE = 0.467$, $z = 4.435$, $p < 0.001$). That is, in both experiments, speakers looked more at the patient when producing a passive utterance. Overall, there were fewer patient fixations in Experiment 2 than in Experiment 1, resulting in a main effect of cue type ($\beta = 2.436$, $SE = 0.42$, $z = -5.806$, $p < 0.001$). An overview of all contrasts is given in Table S.6 in the Supplementary Materials.

Bayesian statistics

Most intriguingly, in Experiment 2, as in Experiment 1 before, we did not find a significant effect of case marking, that is a difference between ambiguously case-marked feminine nouns and unambiguously case-marked masculine nouns. To obtain more information about the uncertainty of our estimates of the case effect, we also analysed our syntactic choice data (Experiment 1 and Experiment 2 combined) with the help of a Bayesian model. For this, we used the R package brms (Bürkner, 2018). We used weakly informative priors to fit the model; the code is available on OSF at https://osf.io/5wkc8/. Table 10 shows the regression coefficients and 95% credible intervals of the model.

Figure 7. Proportion of fixations on patient (coloured) and agent (grey) for ambiguously case-marked feminine nouns (solid lines) and unambiguously case-marked masculine nouns (dashed lines) in Experiment 2. Since for the visualisation of the data we also included other on-screen looks, the proportion of fixations does not add up to one. The dotted vertical line indicates the average speech onset, the two solid vertical lines the analysis window.

Table 8. Mixed-effects logistic regression model on looks to the patient relative to looks on patient and agent for the active voice utterances in Experiment 2.

|       | $\beta$ | $SE$ | $z$-value | $p$-value |
|-------|---------|------|-----------|-----------|
| Intercept (fem./masc.) | $-4.531$ | $0.540$ | $-8.391$ | $<0.001^{***}$ |
| Case | $0.677$ | $0.795$ | $0.852$ | $0.394^{**}$ |

Table 9. Mixed-effects logistic regression model on looks to the patient relative to looks on patient and agent for the passive voice utterances in Experiment 2.

|       | $\beta$ | $SE$ | $z$-value | $p$-value |
|-------|---------|------|-----------|-----------|
| Intercept (fem./masc.) | $-2.327$ | $0.702$ | $-3.313$ | $0.001^{***}$ |
| Case | $-0.634$ | $0.761$ | $-0.833$ | $0.405$ |
the range of $-1$ and $1$. We take this as indication that a case effect, if existent, can only be small.

**Discussion**

In Experiment 2, the cue was not only informative regarding the patient’s visuo-spatial position, like in Experiment 1, but also provided a preview of the patient. This manipulation aimed at making case-related (gender) information available prior to scene onset, pre-activating the patient referent on a conceptual level and, probably, already leading to retrieval of the respective lemma.

With respect to syntactic choice, results for Experiment 2 are in line with those for Experiment 1 since we found no difference between ambiguously case-marked feminine nouns and unambiguously case-marked masculine nouns: the production of non-canonical structures was not modulated by the ambiguity in case marking. A Bayesian model calculated with the joined dataset showed that an effect of case marking did not exist or – if it existed – would only be small. We thus feel safe to conclude that case marking does not play much of a role for the choice of a syntactic structure in German.

Changing the type of cue from a simple perceptual cue to a referential cue increased the production of non-canonical sentences substantially, from 7% in Experiment 1 to 32% in Experiment 2. Like in Experiment 1, non-canonical structures consisted exclusively of passives. Hence, referential cueing had a stronger impact on non-canonical sentence production than simple perceptual cueing but, crucially, led to the same finding as regards the type of non-canonical structure produced. When compared to data from English (Myachykov et al., 2012; Myachykov et al., 2018a), it becomes obvious that speakers of German are differently affected by referential cueing than English speakers. Myachykov et al. (2012) reported no difference between an explicit 700 ms simple cue and an explicit 700 ms referential cue as regards syntactic choices for English native speakers (see also Myachykov et al., 2018a). Note that in addition to an effect of cue type our data also provide some indication for an influence of the position of the referential cue. While we elicited 32% passives for a left-positioned referential cue, Esaulova et al. (2020) elicited 25% passives (26% for left-positioned patients, 23% for right-positioned patients) for a referential cue that was presented in the centre of the screen. Future research may want to determine the role of cue position more systematically.

As regards sentence-planning latencies for Experiment 2, there was again no significant difference between ambiguously case-marked feminine nouns and unambiguously case-marked masculine nouns. Experiment 2 therefore confirms that the lack of a case-marking effect in SOTs is not to be explained by the low number of produced passives observed in Experiment 1. Unlike in Experiment 1, German speakers in Experiment 2 were not faster in producing an active utterance than a passive utterance, suggesting that referential cueing facilitated the production of passives. This finding is in line with Esaulova et al. (2020), who also reported reduced sentence-onset latencies for passives.

Supporting syntactic choice and SOT data, eye-gaze analyses of the 200–600 ms time window for active and passive voice utterances in Experiment 2, like for Experiment 1, showed no difference between ambiguously case-marked feminine nouns and unambiguously case-marked masculine nouns. Thus, we could replicate our findings from Experiment 1 based on a substantially higher number of passive observations (299 trials vs. 69 for Experiment 1). Importantly, the eye-gaze data showed that the cueing manipulation was still effective in guiding speakers’ gaze to patients at scene onset. Like in Experiment 1, the majority of first fixations were on the patient. Hence, the preference for the agent observed in the analysis window in Experiment 2 cannot be attributed to any baseline differences.

The agent preference for passive utterances observed in the eye-gaze pattern in Experiment 2 as opposed to Experiment 1 may seem puzzling at first sight. For the active utterances, the immediate preference for the agent over the patient supports our prior conclusion about linear incremental sentence planning in German. For the passive utterances, in contrast, the direct linkage between eye fixations and word order is missing. Previous research, however, indicates that this observation is systematic when it comes to referential cueing. Eye-tracking analyses in Esaulova et al. (2020) had shown that, before speech onset, participants spent more time looking on agents than patients after referential cueing. However, gaze durations differed between utterance types, with more/longer looks on patients for passives than actives. This is in line with the eye-gaze pattern we observed in Experiment 2 for our pre-defined analysis window. Cueing the patient

| Estimate | Est. Error | l-95% CI | u-95% CI |
|----------|------------|----------|----------|
| Intercept (fem.) | −4.82 | 0.86 | −6.63 | −3.25 |
| Case | 0.14 | 0.65 | −1.18 | 1.40 |
referent activated conceptual information about the referent, so that at scene onset attention was rather dedicated to the new referent, the agent. Nevertheless, we observed more looks on the patient for passive utterances than active utterances and the overall gaze pattern before speech onset resembled the pattern in Experiment 1. This lets us suggest that although cue type had an effect on fixations on patient and agent in an initial time window, it is still highly plausible that the linear incremental strategy was applied to describe transitive events in both experiments.

**General discussion**

In the current study we examined the role of morphological case marking in sentence production after attentional cueing. Previous studies demonstrated that speakers of different languages varied in the extent to which their sentence production was affected by visual cueing (e.g. Hwang & Kaiser, 2015; Myachykov & Garrod, 2008). Whereas speakers of English were strongly affected by visual patient cueing, speakers of German (and other languages) were less inclined to produce passives in the same situation. Here we tested whether differences in case marking could underlie these observed discrepancies. To achieve this goal, we took advantage of the fact that in German subject and object case are unambiguously marked for masculine nouns but not for feminine nouns. Consequently, the current within-language comparison allowed us to specifically study the impact of overt case marking while controlling for other properties that can lead to confounds in cross-linguistic comparisons (e.g. differences in syntactic flexibility).

The most striking finding from our study was that the possibility to ambiguously case mark an entity as subject or object did not affect the description of transitive events in German. We found no difference between masculine nouns and feminine nouns in the production of non-canonical structures. In both experiments, speakers produced passive but not active object-initial sentences with a similar proportion for masculine nouns and feminine nouns (Experiment 1: masc. 8%, fem. 7%; Experiment 2: masc. 33%, fem. 31%). The Bayesian model we calculated in addition to the frequentist model supports our conclusion that an effect of case marking on structural choice is likely to be non-existent in our data. Converging evidence for this conclusion comes from speech onset times and speakers’ eye-gaze behaviour, where we found no differences for masculine nouns and feminine nouns. Especially the latter indicated that sentence planning was unaffected by case marking. Below, we will discuss the theoretical implications of our findings as regards (i) the different effects of patient-cueing on syntactic choice in English versus case-marking languages, (ii) the effect of case marking on sentence-planning strategies and (iii) effects of cue type. Finally, we will discuss the limitations of our study.

**The effect of patient-cueing on syntactic choice**

We introduced two accounts that may explain why the effect of patient-cueing on syntactic choice differs cross-linguistically. First, speakers of case-marking languages may rely to a different extent on grammatical role assignment or linear ordering when accommodating referent accessibility (e.g. Tomlin & Myachykov, 2015). Second, speakers of case-marking languages may be hindered in structure planning by a prominent patient, the more so, the more structural alternatives available (Hwang & Kaiser, 2015). Both accounts have in common that they appoint a key role to how languages assign grammatical functions: by word order or by case marking. In our study, particularly in Experiment 1, speakers demonstrated a strong preference for agents in sentence-initial position. In case they did not choose an agent-before-patient active structure but, instead, placed the patient in sentence-initial position, they assigned the patient the subject role and made use of voice alternation. Thus, German speakers mapped the most prominent entity to the most prominent grammatical role, the subject. It is possible that during production German speakers show a strong subject-first preference because, as assumed by the “mapping mechanism account”, grammatical role assignment is prioritised over linear ordering. Yet, the account does not explain why the effect of patient-cueing in our study was reduced when compared to English. While in Experiment 1 we elicited 7% passives after explicit patient-cueing, Myachykov et al. (2018a) elicited around 25% passives after explicit patient-cueing in English. Hence, the “mapping mechanism account” alone does not sufficiently explain the observed cross-linguistic differences.

What about the “syntactic flexibility account”? Does the number of structural alternatives reduce the effect of patient-cueing on syntactic choice? If we leave aside cleft constructions, English speakers can choose from two structures (active SVO or passive), whereas German speakers can choose from three structures (active SVO, active OVS, passive). The availability of an additional syntactic structure in German, active OVS, may affect sentence planning in a way that German speakers resort to the canonical active SVO structure and produce fewer passives than English speakers.
Following this logic, Russian speakers, who can choose from a total of 12 structures (Myachykov et al., 2013), should be even less affected by patient-cueing than German and English speakers. However, this was not the case. Instead, in Experiment 1 we elicited 7% patient-initial sentences (passives) after an explicit patient cue, which is much lower than the 22% patient-initial sentences (mostly object-initial actives) that were obtained by Pokhoday et al. (2019) in a similar experimental condition tested with Russian speakers. Hence, neither of the two accounts can fully explain the data.

What could explain the two findings that German speakers are less influenced by visual cueing than speakers of English and Russian, and produce passives rather than active OVS? While in English grammatical function assignment is mostly realised through word order, that is, grammatical function is encoded positionally, German can encode certain functions both ways, positionally and by case marking (e.g. Levelt, 1989). Hence, German speakers may prefer to assign a prominent patient the default case (nominative), marking the patient as the subject, which results in a passive sentence where the subject is placed in sentence-initial position. As further discussed below, German speakers may do so not only to keep both mapping mechanisms, grammatical role assignment and linear ordering, in alignment, but because they strive to avoid conflicting information (first NP is not the nominative marked subject). In contrast, Russian speakers have been shown to avoid the dispreferred passive structure and, rather than using the dominant grammatical role assignment mechanism, prioritised the linear ordering mechanism to accommodate a prominent patient (e.g. Tomlin & Myachykov, 2015). We may speculate that they do so especially because case marking is such a reliable cue to encode and identify grammatical functions and thematic roles in Russian. Unlike in German, ambiguity in case marking is rare in Russian (Myachykov, 2007, p. 111). Hence, there is no need for speakers to rely on word order. For example, Kempe (1999) found that during comprehension German speakers relied more on word order than case to identify the agent in a transitive event in contrast to Russian speakers. Maybe it was the ambiguity in case marking that prevented German speakers from producing active OVS sentences in the current study. Recall that due to case syncretism, for feminine nouns speakers could start the utterance with the cued feminine patient and continue with a patient-initial active or a passive sentence. Yet, when speakers produced a non-canonical structure, they chose a patient-initial passive sentence. We might speculate that our speakers took a potential listener into account and avoided OVS utterances that could be easily misinterpreted as active SVO sentences. Schouwenaars et al. (2018), for example, showed that German-speaking adults and children between seven and ten years already considered the listener when producing which questions. They produced more object questions than passive questions if case cues were available, whereas they produced more passive questions if only agreement cues were available. However, this provides no explanation for the complete absence of object-initial active sentences for masculine nouns in German and Russian.

While future studies could target these open issues, both the mapping mechanism account and the syntactic flexibility account as they stand cannot fully explain why the effect of patient-cueing on structural choice differs cross-linguistically. However, a more differentiated model, combining assumptions of both accounts with language-specific information regarding validity of case-marking cues to identify the agent in a sentence, might be able to capture the difference between English and German as well as between German and Russian.

Nevertheless, our central finding remains that it did not matter for the choice of a syntactic structure whether case marking was overt and unambiguous or not. Even when German speakers produced sentences with feminine NPs – devoid of unambiguous morphological case marking and hence more analogous to English noun phrases – German speakers did not produce a comparable proportion of passives as English speakers when the patient was in the spotlight of attention. Thus, our findings indicate that morphological case marking is not the main property driving the observed discrepancies between English and German. Whatever the cause for the cross-linguistic variation as regards effects of attention manipulation on syntactic choices, our findings suggest that cross-linguistic differences cannot be simply attributed to case marking alone.

The effect of case marking on sentence-planning strategies

Researchers have suggested that the necessity of early morphological marking on sentence constituents
might affect sentence planning (e.g. Norcliffe et al., 2015). Therefore, we did not only analyse sentence-planning latencies in our study but also speakers’ eye-gaze behaviour within an initial time window that has been associated with conceptualisation. Recall that early morphological encoding has been assumed to invoke hierarchical planning (e.g. Norcliffe & Konopka, 2015), which, in eye-tracking, should be reflected in a similar distribution of looks to agent and patient. Against this background, our study addressed the question whether planning may be hierarchical when case is marked overtly on the first NP and incremental when not overtly marked.

In our study, we found no indication for hierarchical incremental planning when two masculine nouns were present in the event scene. Masculine nouns in German have distinctive case marking on NPs, hence the speaker’s syntactic choice is set as soon as the sentence-initial NP is uttered. Instead, for both masculine nouns and feminine nouns we found indication for linear incremental planning, in line with previous findings for English speakers in Gleitman et al. (2007).

In Experiment 1 and Experiment 2, speakers showed an immediate preference for the agent over the patient when planning an active SVO sentence. In contrast, speakers showed a preference for the patient when planning a passive sentence, although the eye-gaze behaviour was affected by the type of cue, which will be further discussed below. The eye-gaze results align with the results for sentence-planning latencies, which showed no difference between masculine nouns and feminine nouns either.

We conclude from these findings that overt and unambiguous case marking on NPs does not lead to hierarchical incremental planning, at least not in German speakers when tested in the current set-up. German speakers start sentence planning with the first increment that is to become the nominative-marked sentential subject, irrespective of the nouns’ grammatical gender. Future research might want to investigate which sentence-planning strategy is applied by Russian speakers, who, as discussed above, when not using a canonical active structure tend to place non-nominative marked objects in sentence-initial position to accommodate a prominent patient.

**Effects of cue type**

In the current study we used two different cue types: a simple perceptual cue in Experiment 1, a red dot presented at the position where the patient was to appear in the event scene, and a referential cue in Experiment 2, a picture of the patient at the same position where the patient was to appear in the event scene. Cue type had an effect on all measures: syntactic choice, sentence-planning latency as measured by speech onset times and eye gaze. We elicited more passives when a referential cue preceded the event scene as compared to a simple perceptual cue. Moreover, speakers of German were faster to produce a passive sentence after referential cueing than after simple cueing. Finally, looks to the patient and agent between 200 and 600 ms after scene onset were affected by the type of cue, with fewer looks to patients after referential cueing.

The effect of cue type in our study differs quite remarkably from previous findings for English (Myachykov et al., 2012; Myachykov et al., 2018a), indicating that German speakers are not only differently affected by patient-cueing than English speakers as seen for the proportion of passives produced, but are also differently affected by the type of cue. Unlike English speakers, German speakers’ production of passives seemed to be boosted by referential cueing and, thus, additional conceptual priming: When speakers’ attention was directed towards the patient by a cue that was also informative in terms of lexical information, German speakers produced more passives and were faster in doing so. The eye-gaze data for Experiment 2 indicated that as the patient was activated by the referent preview, attention in terms of looks to the patient was no longer required. Crucially, however, the type of cue did not affect overall sentence planning. That is, irrespective of whether the cue tapped into an early planning stage (simple cue) or a somewhat later stage with the patient being already conceptualised and probably a respective lexical entry accessed (referential cue), German speakers tended to focus on the referent to be placed in sentence-initial position before starting the utterance.

**Limitations**

In this paper, we aimed to answer how case marking in German affects the description of transitive events against the background of cross-linguistic differences. Ideally, the identification of cross-linguistic differences would draw on a design that uses the same materials and methodology (e.g. same cue type, same cue length) across languages. At present, the variability of existing studies limits such cross-linguistic comparisons. In the current study, we explored the influence of case marking in a within-language design. Future studies, however, should target this issue in a between-language design, employing the same materials and methodology.

Previous studies also reported within-language variability as regards the use of different sentence-planning strategies (e.g. Norcliffe et al., 2015). Although we
found no indication of different strategies for unambiguously case-marked masculine nouns and ambiguously case-marked feminine nouns in our study, this does not necessarily mean that German speakers apply a linear incremental planning strategy across the board. As pointed out to us by an anonymous reviewer, we cannot rule out that the current design enforced a linear incremental planning strategy in Experiment 2 since speakers had sufficient time to retrieve the name of the patient referent during the preview. This means that they could start their utterance with the retrieved noun and then proceed increment by increment. Follow-up experiments could, hence, change the procedure to minimise the preview time. While this limitation might hold for Experiment 2, it does not account for the same findings in Experiment 1, where no referents were previewed prior to scene onset. The findings from Experiment 1, nevertheless, indicated that German native speakers made use of a linear incremental strategy when describing a simple transitive event. Crucially, even if participants were biased towards a linear incremental planning strategy, we do not think that this renders our main finding – namely the absence of (gradual) differences for unambiguous versus ambiguous case marking across measures – inadequate.

Conclusion

The current study showed that there was no difference between ambiguous and unambiguous case marking neither on syntactic choices nor on sentence planning. Therefore, our findings demonstrate that case marking per se does not underlie previously observed discrepancies between English and German speakers concerning sentence production during visual cueing. Overall, German speakers displayed a strong preference for agent-initial active sentences to describe transitive events, even if the prominence of the patient had been increased by visual cueing. More patient-initial sentences could be elicited by simultaneous perceptual and conceptual priming via a referential cue. When German speakers produced a non-canonical structure, they made use of voice alternation but not object topicalization, irrespective of ambiguity in case marking. Cross-linguistic variation in the effect of patient-cueing on syntactic choices seems to be best explained by a combination of language-specific properties such as reliance on syntactic mapping mechanisms, the availability of alternative syntactic structures and, presumably, the reliance on case marking cues to identify the subject/agent. Unlike what had been proposed previously, we found no indication that morphological case marking on NPs led to hierarchical incremental sentence planning. Instead, German speakers in our study used a linear incremental planning strategy, as shown by the eye-tracking data, for both ambiguously case-marked feminine nouns and unambiguously case-marked masculine nouns. The recording of speakers’ eye gaze during event descriptions therefore provides a fruitful avenue for the investigation of language production, illuminating the interplay between sentence planning strategies and language-specific properties.

Notes

1. Why the voice-based alternation is less available in Russian is not yet clear and might have several reasons, including the infrequency of passives. As suggested by Myachykov, syntactic alternations in general are less productive in Russian than in English.
2. As pointed out to us by an anonymous reviewer, it should be noted that the authors excluded all responses that, contra the instruction provided to the speakers, did not mention both characters (in total, 7% of the trials were excluded). Since the most frequent word order in Korean is OV with a null subject (Kim, 2008), both the instruction and coding protocol are problematic and render the comparison difficult.
3. In previous studies, time windows were defined rather arbitrarily. Our assumptions about when and how the two planning strategies should surface in eye-gaze behaviour is based on Konopka et al. (2018).
4. Word form frequencies were collected using the TAGGED-M öffentlich corpus from the Archive of written language, Institute for German Language, Mannheim, Germany. This morphologically annotated corpus allowed us to specify the word class (i.e. noun) and the number (i.e. singular) of the searched terms in order to exclude irrelevant items (e.g. “braut” as a verb (“brews”) when searching for “Braut” “Bride”) and distinguish between same-form words (e.g. “Piraten” as in plural “pirates” as opposed to the accusative of singular “pirate”).
5. The Miles test was used to determine the dominant eye. Participants extended both arms and created an opening with their hands through which they fixated a target point on the wall. When instructed to move their hands slowly towards their face while fixing the point, the opening was drawn towards the dominant eye.
6. We did not exclude sentences in which participants used a neuter noun (e.g. das Mädchen – “the girl”) instead of a feminine noun because feminine and neuter nouns are both ambiguous between nominative and accusative case.
7. For the logistic regression, counts of events in each group were used rather than raw binary values. By aggregating the binary values of a time bin into a count one can ease computation through reducing the size of the dataset without losing too much information (arguably the resolution by which fixations change is even larger than 20 ms), while at the same time reducing the correlation between neighbouring observations, that is, reducing autocorrelation of the residual; see the problem of eye-movement-based
dependencies in Barr (2008). The reader interested in the specific formula is referred to the script and documentation on OSF.

8. The percentage of passives in Experiment 1 was slightly lower than in Esaulova et al. (2021). A possible explanation for this is the difference in the design, with a patient cue preceding each and not only some experimental items.

9. The eye-movement pattern for the active utterances allows an alternative interpretation, namely initial event apprehension within the first 200–300 ms followed by fixations on characters in their order of mention in line with Griffin and Bock (2000). However, we question that the participants processed the relationship between patient and agent while initiating a saccade towards the agent, but rather attribute this shift of attention to our cueing manipulation.

10. A comparison with the Finnish data from Myachykova et al. (2011a) is difficult because the authors used an implicit (less than 80 ms) patient cue.

11. In canonically ordered active sentences, the first NP is the subject and agent, whereas for passive sentences the first NP is the subject but not the agent. The most conflicting information is thus provided with an object-initial active sentence (first NP is the accusative object and patient).

Acknowledgements

We are thankful to our colleagues at the Collaborative Research Center 1252 "Prominence in Language" for their feedback and support on this project. We would also like to thank Cecilia Puebla Antunes for her help in the preparation of the picture stimuli, Sabine Reuters and Elysia Seidel for their help in the preparation of the data, as well as all the assistants who helped with the data collection.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Deutsche Forschungsgemeinschaft: [Grant Number Project-ID 281511265 SFB 1252].

ORCID

Judith Schlenter http://orcid.org/0000-0001-8676-6629

References

Bader, M., & Häussler, J. (2010). Word order in German: A corpus study. *Lingua*, 120(3), 717–762. https://doi.org/10.1016/j.lingua.2009.05.007

Bakker, D. (1998). Flexibility and consistency in word order patterns in the languages of Europe. In A. Siewierska (Ed.), *Constituent order in the languages of Europe* (pp. 383–420). Mouton de Gruyter.

Barr, D. J. (2008). Analyzing ‘visual world’ eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Blake, B. J. (2001). *Case*. Cambridge University Press.

Bock, K., Irwin, D. E., & Davidson, D. J. (2004). Putting first things first. In J. M. Henderson, & F. Ferreira (Eds.), *The interface of language, vision and action: Eye movements and the visual world* (pp. 249–278). Psychology Press.

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99(1), 150–171. https://doi.org/10.1037/0033-295X.99.1.150

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5(9/10), 341–345.

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2002). Grammar overrides frequency: Evidence from the online processing of flexible word order. *Cognition*, 85(2), B21–B30. https://doi.org/10.1016/S0010-0272(02)00076-8

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass*, 3(1), 19–58. https://doi.org/10.1111/j.1749-818X.2008.00099.x

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, 26(2), 211–252.

Burmester, J., Spalek, K., & Wartenburger, I. (2019). Visual attention-capture cue in depicted scenes fails to modulate online sentence processing. *Dialogue & Discourse*, 10(2), 79–104.

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Esaulova, Y., Dolscheid, S., & Penke, M. (2021). All it takes to produce passives in German. In V. Torrens (Ed.), *Syntax processing* (pp. 75–107). Cambridge Scholars Publishing.

Esaulova, Y., Penke, M., & Dolscheid, S. (2019). Describing events: Changes in eye movements and language production due to visual and conceptual properties of scenes. *Frontiers in Psychology*, 10, 1–15. https://doi.org/10.3389/fpsyg.2019.00835

Fedzechkina, M., Newport, E. L., & Jaeger, T. F. (2017). Balancing effort and information transmission during language acquisition: Evidence from word order and case marking. *Cognitive Science*, 41, 416–446. https://doi.org/10.1111/cogs.12346

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and
utterance formulation. *Journal of Memory and Language*, 57 (4), 544–569. https://doi.org/10.1016/j.jmla.2007.01.007

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11(4), 274–279. https://doi.org/10.1111/1467-9280.00255

Hwang, H., & Kaiser, E. (2014). Having a syntactic choice is not always better: The effects of syntactic flexibility on Korean production. *Language, Cognition and Neuroscience*, 29(9), 1115–1131. https://doi.org/10.1080/23273798.2013.875212

Hwang, H., & Kaiser, E. (2015). Accessibility effects on production vary cross-linguistically: Evidence from English and Korean. *Journal of Memory and Language*, 84, 190–204. https://doi.org/10.1016/j.jmla.2015.06.004

Kempe, V. (1999). Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes*, 14 (2), 129–171. https://doi.org/10.1080/016909699386329

Kim, T. (2008). Subject and object markings in conversational Korean [Doctoral dissertation]. State University of New York at Buffalo. Available from ProQuest Dissertations & Theses Global. (304369360). Retrieved from https://www.proquest.com/dissertationstheses/subject-object-markings-conversational-korean/docview/304369360/se-2?accountid=10218

Konopka, A. E., Meyer, A., & Forest, T. A. (2018). Planning to speak in L1 and L2. *Cognitive Psychology*, 102, 72–104. https://doi.org/10.1016/j.cogpsych.2017.12.003

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). LmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). Package ‘emmeans’ (Version 1.4.8). https://github.com/rleveland/emmeans

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, 53(4), 372–380. https://doi.org/10.3758/BF03206780

Matuschek, H., Klieg, R., Vasisht, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. https://doi.org/10.1016/j.jmla.2017.01.001

Myachykov, A. (2007). Integrating perceptual, semantic, and syntactic information in sentence production [Doctoral dissertation]. University of Glasgow. Available from https://www.researchgate.net/publication/33053551_Integrating_perceptual_semantic_and_syntactic_information_in_sentence_production

Myachykov, A., & Garrod, S. (2008). Perception and word order in Russian and Finnish sentence production. In W. Ramm, & C. Fabricius-Hansen (Eds.), *Lineairisation and segmentation in discourse: Multidisciplinary approaches to Discourse (MAD 08)* (pp. 87–98). University of Oslo.

Myachykov, A., Garrod, S., & Scheepers, C. (2011). Perceptual priming of structural choice during English and Finnish sentence production. In R. K. Mishra, & N. Srinivasan (Eds.), *Language-cognition interface: State of the art* (pp. 53–71). LINCOM Europa.

Myachykov, A., Garrod, S., & Scheepers, C. (2011a). Perceptual priming of structural choice during English and Finnish sentence production. In R. K. Mishra, & N. Srinivasan (Eds.), *Language-cognition interface: State of the art* (pp. 53–71). LINCOM Europa.

Myachykov, A., Garrod, S., & Scheepers, C. (2012). Attention and memory play different roles in syntactic choice during sentence production. *Discourse Processes*, 55(2), 218–229. https://doi.org/10.1080/0163853X.2017.1330044

Myachykov, A., Pokhoday, M., & Tomlin, R. (2018b). Attention and structural choice in sentence production. In S.-A. Rueschemeyer & G. Gaskell (Eds.), *Oxford handbook of psycholinguistics* (pp. 529–547). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198786825.013.22

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: The case of English and Russian. *Quarterly Journal of Experimental Psychology*, 66(8), 1601–1619. https://doi.org/10.1080/17470218.2012.754910

Myachykov, A., Thompson, D., Garrod, S., & Scheepers, C. (2012). Referential and visual cues to structural choice in visually situated sentence production. *Frontiers in Psychology*, 2, 1–9. https://doi.org/10.3389/fpsyg.2011.00396

Myachykov, A., Thompson, D., Scheepers, C., & Garrod, S. (2011b). Visual attention and structural choice in sentence production across languages. *Language and Linguistics Compass*, 5(2), 95–107. https://doi.org/10.1111/j.1749-818X.2010.00265.x

Norcliffe, E., & Konopka, A. E. (2015). Vision and language in cross-linguistic research on sentence production. In R. K. Mishra, N. Srinivasan, & F. Huettig (Eds.), *Attention and vision in language processing* (pp. 77–96). Springer. https://doi.org/10.1007/978-81-322-2443-3_5

Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, 30(9), 1187–1208. https://doi.org/10.1080/23273798.2015.1006238

Osborne, J. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research, and Evaluation*, 15, 1–9. https://doi.org/10.7275/QBPC-GK17

Pokhoday, M., Shtyoer, Y., & Myachykov, A. (2019). Effects of visual priming and event orientation on word order choice in Russian sentence production. *Frontiers in Psychology*, 10, 1–8. https://doi.org/10.3389/fpsyg.2019.01661

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32*(1), 3–25. https://doi.org/10.1080/01670269084048231

Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, 42(2), 168–182. https://doi.org/10.1006/jmla.1999.2668

Primus, B. (1998). The relative order of recipient and patient in the languages of Europe. In A. Siewierska (Ed.), *The languages of Europe* (pp. 421–474). Mouton de Gruyter.

R Core Team. (2020). R: A language and environment for statistical computing. https://www.R-project.org/

Sauppe, S. (2017a). Symmetrical and asymmetrical voice systems and processing load: Pupillometric evidence from sentence production in Tagalog and German. *Language*, 93(2), 288–313. https://doi.org/10.1353/lan.2017.0015

Sauppe, S. (2017b). Word order and voice influence the timing of verb planning in German sentence production. *Frontiers
Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin, R. D. J., & Levinson, S. C. (2013). Dependencies first: Eye-tracking evidence from sentence production in Tagalog. In M. Knauff, M. Pauen, N. Sebanz, & E. Wachsmuth (Eds.), Proceedings of the 35th annual meeting of the Cognitive Science society (pp. 1265–1270). Cognitive Science Society.

Schouwenaars, A., Hendriks, P., & Ruigendijk, E. (2018). German children’s processing of morphosyntactic cues in wh-questions. Applied Psycholinguistics, 39(6), 1279–1318. https://doi.org/10.1017/S0142716418000334

Siewierska, A. (1998). Variation in major constituent order. In A. Siewierska (Ed.), Constituent order in the languages of Europe (pp. 475–552). Mouton de Gruyter.

Suitner, C., Maass, A., Navarrete, E., Formanowicz, M., Bratanova, B., Cervone, C., Hakoköngäs, J. E., Kuppers, T., Lipourli, E., Rakić, T., Scatolon, A., Teixeira, C. P., Wang, Z., Sobral, M. P., & Carrier, A. (2021). Spatial agency bias and word order flexibility: A comparison of 14 European languages. Applied Psycholinguistics, 42(3), 657–671. https://doi.org/10.1017/S0142716420000831

Tanenhaus, M. K., & Trueswell, J. C. (2006). Eye movements and spoken language comprehension. In M. J. Traxler, & M. A. Gernsbacher (Eds.), Handbook of Psycholinguistics (2nd ed., pp. 863–900). Elsevier/Academic Press.

Tomlin, R., & Myachykov, A. (2015). Attention and salience. In E. Dąbrowska, & D. Divjak (Eds.), Handbook of Cognitive Linguistics (pp. 31–52). De Gruyter Mouton.

van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing “easy” and “hard” events with preferred and dispreferred syntactic structures. Journal of Memory and Language, 71(1), 124–144. https://doi.org/10.1016/j.jml.2013.11.001