Juliane Burmester | Antje Sauermann | Katharina Spalek | Isabell Wartenburger

Sensitivity to salience

linguistic vs. visual cues affect sentence processing and pronoun resolution

Suggested citation referring to the original publication:
Language, Cognition and Neuroscience 33 (2018) 6, 784–801
DOI http://dx.doi.org/10.1080/23273798.2018.1428758
ISSN (print) 2327-3798
ISSN (online) 2327-3801

LANGUAGE, COGNITION AND NEUROSCIENCE, 2018
VOL. 33, NO. 6, 784–801
https://doi.org/10.1080/23273798.2018.1428758

Routledge
Taylor & Francis Group

REGULAR ARTICLE

OPEN ACCESS Check for updates

Sensitivity to salience: linguistic vs. visual cues affect sentence processing and pronoun resolution

Juliane Burmester [a], Antje Sauermann[b], Katharina Spalek[c] and Isabell Wartenburger[a]

[a]Department of Linguistics, Cognitive Sciences, University of Potsdam, Potsdam, Germany; [b]Centre for General Linguistics (ZAS), Berlin, Germany; [c]Department of German Language and Linguistics, Humboldt-Universität Berlin, Berlin, Germany

ABSTRACT

Sentence comprehension is optimised by indicating entities as salient through linguistic (i.e., information-structural) or visual means. We compare how salience of a depicted referent due to a linguistic (i.e., topic status) or visual cue (i.e., a virtual person's gaze shift) modulates sentence comprehension in German. We investigated processing of sentences with varying word order and pronoun resolution by means of self-paced reading and an antecedent choice task, respectively. Our results show that linguistic as well as visual salience cues immediately speeded up reading times of sentences mentioning the salient referent first. In contrast, for pronoun resolution, linguistic and visual cues modulated antecedent choice preferences less congruently. In sum, our findings speak in favour of a significant impact of linguistic and visual salience cues on sentence comprehension, substantiating that salient information delivered via language as well as the visual environment is integrated in the current mental representation of the discourse.

ARTICLE HISTORY
Received 4 November 2016
Accepted 5 December 2017

KEYWORDS
Topic status; eye gaze; visual context; reading times; antecedent choice

1. Introduction

Human communication often takes place in form of both verbal and non-verbal behaviour which is adjusted to the situational needs and communicative objectives of the interlocutors. As for the verbal (linguistic) modality, speakers use information-structural means (e.g., word order, referring expressions) to adapt their utterances to the needs of their listeners and to convey the intended meaning most properly and effectively (Ariel, 1988; Gundel, Hedberg, & Zacharski, 1993; Halliday, 1967). For instance, speakers typically place the *topic* (i.e., that part of information about which the speaker intends to increase the listeners' knowledge; Gundel, 1985) in prominent sentence-initial position to induce it as salient and to facilitate listeners' processing (e.g., MacWhinney, 1977). As for the non-verbal (visual) modality, speakers typically look at what they refer to and/or use co-referential gestures that induce entities as salient to improve listeners' comprehension and to support joint attention mechanisms (Baldwin, 1995; Goodrich Smith & Hudson Kam, 2015; Staudte, Crocker, Heloir, & Kipp, 2014). Hence, listeners' challenge is to simultaneously link the linguistic input to the situational (visual) environment and to pay attention to coherence relations between both, in order to draw inferences and generate expectations about the upcoming discourse (Kaiser & Trueswell, 2004; Kamide, Altmann, & Haywood, 2003). To optimise comprehension, listeners might profit from information that is salient in the linguistic and/or visual context. Our study addresses the question how a linguistic vs. visual salience cue affects different aspects of sentence comprehension. First, we wanted to see how the initial stages of **sentence processing** were affected. Differences in initial processing are most likely reflected in reading times, with shorter reading times indicating easier processing. Second, we also investigated if and how the cues influenced later, more interpretative stages of comprehension. For this, we looked at pronoun processing. In order to process a pronoun correctly, comprehenders need to identify its antecedent in the preceding discourse. Different types of cues might make this process easier or harder or might even change which discourse element is chosen as antecedent. Therefore, we looked at later comprehension processes through **pronoun resolution** (both its ease, as measured in reading times, and the eventual antecedent choice).

Whereas many studies demonstrate a significant impact of exclusively linguistically presented information on sentence processing and pronoun resolution (e.g.,

CONTACT Juliane Burmester juliane.burmester@uni-potsdam.de
Supplemental data for this article can be accessed at https://doi.org/10.1080/23273798.2018.1428758.

Bornkessel & Schlesewsky, 2006; Cowles, Walenski, & Kluender, 2007), a growing branch of recent psycholinguistic research supports the close relation of linguistic and visual information during sentence and pronoun comprehension (e.g., Arnold & Lao, 2015; Crocker, Knoeferle, & Mayberry, 2010) as well as production (e.g., Fukumura, van Gompel, & Pickering, 2010; Vogels, Krahmer, & Maes, 2013). The theory of *Mental Models* (e.g., Johnson-Laird, 1980) and other discourse models (e.g., Bower & Morrow, 1990; Dik, 1997) have already assembled the idea that interlocutors (each) build a mental representation of the discourse which is dynamically updated based on linguistic input as well as the whole situation. Mental representations are understood as non-linguistic meaning representations, influenced by memory, other mental processes, and world knowledge. Different seminal frameworks link attention state and memory to a mental representation's accessibility (or cognitive status) (e.g., *Accessibility Theory* by Ariel, 1988; *Givenness Hierarchy* by Gundel et al., 1993; *Structure Building Framework* by Gernsbacher, 1991; see Arnold & Lao, 2015 for the different views about how attention might be involved in the accessibility of discourse information). These different frameworks offer theoretical explanations about how differences in accessibility are reflected in the syntactic structure and/or choice of the referential expression. The sentence-initial position is attributed to high accessibility as it is easily paid attention to, enhancing information retrieval for the listener (Gernsbacher, 1991; Levelt, 1989; MacWhinney, 1977). In addition, reduced referential forms (e.g., pronouns) typically refer to highly accessible information (e.g., topic) (e.g., Ariel, 2001). For the purpose of our study, the term *salience* is used to describe information (i.e., referents) currently in the focus of attention of the addressee. We modulate salience experimentally by linguistic or visual cues that are thought to induce a referent as highly accessible relative to other referents in the mental model (e.g., Burkhardt & Roehm, 2007; Kaiser, 2006).[1]

The present study addresses the question how the modality of a cue that indicates a referent as either linguistically or visually salient affects sentence comprehension in German. Our aim is to better understand whether and how discourse processing, in particular the processing of a discourse referent's information status (i.e., topic status), is grounded in the linguistic domain or if non-linguistic (i.e., visual) cues could affect sentence processing and pronoun resolution similarly. In section 1.1 and 1.2 we will review previous evidence showing ample parallels in the processing of linguistic and visual context information. The evidence supports the following line of argumentation: Salient referents (both when salience is indicated by linguistic or by visual cues) have at least two common information packaging preferences: 1) Regarding sentence processing, salient referents are (expected to be) mentioned first (i.e., in sentence-initial position), and 2) regarding pronoun resolution, salient referents tend to be interpreted as the antecedent referent of a following pronoun. However, until now no study directly compared the impact of cue modality on a discourse referent's accessibility degree by means of the very same experimental material and paradigm. Therefore, we compared if and to what extent salience of a referent due to either a linguistic cue (Exp. 1: topic status) or visual cue (Exp. 2: gaze shift of a virtual person) differ in their capacity to affect initial sentence processing and later pronoun resolution.

Multiple pieces of evidence concerning the impact of linguistic and visual context information are based on the comprehension of German canonical subject-verb-object (SO) vs. non-canonical object-verb-subject (OS) sentences. Due to the strong subject-first preference, the SO order is frequent and easy to process even without context information. In contrast, the OS order is much less frequent and its processing is enhanced if presented in a suitable linguistic or visual context (e.g., object is topic, Bader & Häussler, 2010; Burmester, Spalek, & Wartenburger, 2014; Knoeferle & Kreysa, 2012). Hence, to shed light on the impact of cue modality, it is beneficial to investigate German sentence processing as the modulation of context information should be reflected during sentence processing, especially during processing of the context-sensitive OS word order. Next, we give an overview about how sentence processing and pronoun resolution are affected by accessibility modulations in linguistic and visual contexts.

1.1. The impact of linguistic context on sentence processing and pronoun resolution

In linguistic research, the information status of discourse referents has been proposed to be represented in terms of accessibility degrees or activation states (e.g., for a review see Arnold, 1998; Lambrecht, 1994). The *topic* – what the sentence is about (Reinhart, 1981)– is accompanied with a high degree of mental accessibility (e.g., Ariel, 1988; Givón, 1983). Moreover, topic has been understood as a cognitive rather than formal linguistic concept as it activates the listener's mental representation right at sentence beginning (e.g., Portner, 2007). In German main clauses, which exhibit a strong subject-first preference (e.g., Hemforth, 1993), topics and contrastive entities[2] occur sentence-initially (e.g.,

Frey, 2004; Rosengren, 1993; Speyer, 2008). In fact, topic status and contrastive information can ease processing of non-canonical sentences in German (Burmester et al., 2014; Weskott, Hoernig, Fanselow, & Kliegl, 2011). For the listener, topic accessibility also correlates with a referent's predictability in discourse, which is for instance affected by availability of thematic role information (i.e., knowing that a referent is the agent or patient of the action). In addition to topic status and first-mention, features such as grammatical subject status, animacy, and agent status each contribute to a referent's high accessibility (e.g., Ariel, 1988; Grewe et al., 2006; Jackendoff, 1990; Kaiser, 2011). Accordingly, if all features coincide in the same referent, accessibility might add up to a very high degree. By contrast, if not coincided, different referents might compete for accessibility. Speakers tend to mention highly accessible referents early in the sentence. This first-mention tendency explains in parts when a speaker prefers passive voice over active voice, namely in exactly those cases where, by the use of the passive, the speaker can mention a more accessible referent before a less accessible one (e.g., Flores d'Arcais, 1975; Prat-Sala, 2000; Tannenbaum & Williams, 1968). Hence, during sentence processing the impact of topic status on referent accessibility might interact with first-mention, grammatical role (subject status) and thematic role (agent status) which therefore need to be taken into account in studies investigating how a referent's salience affects sentence processing.

Similar factors seem to impact pronoun resolution: Garnham (2001) and others (e.g., Cowles et al., 2007; Stevenson, Crawley, & Kleinman, 1994) strongly attribute pronoun (and in the broader sense, anaphora) resolution to a referent's accessibility within the mental model. It is generally assumed that the most accessible referent is the most likely antecedent for a following pronoun. For instance, the personal pronoun "*he*" refers to a highly accessible antecedent that has masculine gender and singular number (for a review see Arnold, 1998) and commonly holds a parallel thematic role and/or grammatical function (e.g., Stevenson, Nelson, & Stenning, 1995). In addition, the following features of the antecedent lead to a high accessibility degree: syntactic prominence (i.e., subject status) (e.g., Frederiksen, 1981; Järvikivi, van Gompel, Hyönä, & Bertram, 2005; for German: Bosch, Katz, & Umbach, 2007; Bouma & Hopp, 2007), first-mention (e.g., Järvikivi et al., 2005), discourse prominence (e.g., topic status; Ariel, 1988; Bosch & Umbach, 2007; Colonna, Schimke, & Hemforth, 2012; Cowles, 2007; Rohde & Kehler, 2014) and referential continuity in discourse (e.g., topic status; Frederiksen, 1981; Givón, 1983). Subject status seems to evoke a higher accessibility relative to topic status as revealed by the likelihood of first-mention (Cowles & Ferreira, 2012) or antecedent interpretations (Bouma & Hopp, 2007; Colonna et al., 2012) which might be due to the strong preference of subjects being the (default) topic (at least in English and German). German sentence comprehension is an interesting testbed as it allows disentangling effects of topic (or salience) status on first-mention independent of subject status. This is possible due to the relatively flexible German word order in which reordering of sentence constituents does not change their grammatical function (i.e., SO vs. OS word order).

1.2. The impact of visual context on sentence processing and pronoun resolution

Ariel (1988) already suggested that the degree of mental accessibility is enhanced if the relevant information is attended to in the physical surroundings. Amongst other principles, Osgood and Bock (1977) proposed that information is primarily salient due to perception-based experience which speakers mirror in the natural linear order within a sentence. That is, the constituent which is perceived as most salient is attributed to the leftward sentence position. In addition, with regard to processing of visual information during exposure with a display, it has been found that eye fixations follow a certain pattern: The left located picture strongly tends to be fixated first (Dahan, Tanenhaus, & Salverda, 2007).

The close relation of visual and linguistic information is supported by more recent studies showing that various types of visual cues (e.g., depicted referents in action, a human's or a robot's speaker gaze, or gestures) incrementally affect sentence processing and pronoun resolution (e.g., Crocker et al., 2010 and references therein; Holle et al., 2012; Nappa & Arnold, 2014; Staudte & Crocker, 2011). For instance, processing of German SO and OS sentences (i.e., correct thematic role assignment) is improved by visually depicting the sentential referents in action (i.e., a scene depicting who is performing the action (i.e., agent) with whom (i.e., patient); Knoeferle, Crocker, Scheepers, & Pickering, 2005; Knoeferle, Habets, Crocker, & Münte, 2007). Moreover, processing of German SO and OS sentences is improved by visually depicting the sentential referents *without* depicting the action, but instead modulating listeners' attention by means of visually depicting speaker's gaze (plus head movement) to these referents (Knoeferle & Kreysa, 2012). Concerning pronoun resolution, the preference to interpret pronouns in favour of the first-mentioned referent is modulated by for instance speaker's gaze plus a pointing gesture or an exogenous visual-attentive cue at story onset that draws listeners'

attention to a depicted referent (Arnold & Lao, 2015; Nappa & Arnold, 2014).

Equivalent to the preference that linguistically salient (i.e., topic) referents are (expected) to be mentioned sentence-initially, there is multiple evidence on the production side demonstrating that *visually* salient referents are also preferably mentioned sentence-initially (at least in English). Speakers' eye-movements during sentence production substantiate that what speakers look at in depicted actions is what they mention first: Visual-attentive cues, even when the speaker is unaware of them, indicate a referent as more salient amongst others, and hence as earlier accessible and most likely to be mentioned first (e.g., Gleitman, January, Nappa, & Trueswell, 2007; Myachykov, Thompson, Garrod, & Scheepers, 2012; Tomlin, 1997; for a discussion of broader cognitive and communicative factors that affect first-mention see Bock, Irwin, & Davidson, 2004). In contrast to mental accessibility (e.g., Ariel, 1988), these production studies establish accessibility at the lemma/conceptual level (Bock & Warren, 1985).[3] To the best of our knowledge, only one study directly compared the impact of visual vs. linguistic salience, namely on sentence production: Vogels et al. (2013) found that visually salient (i.e., foregrounded) referents are more likely to be mentioned first (as the subject) but –in contrast to linguistically salient referents– less likely to be referred to with reduced referential expressions (e.g., pronouns). Moreover, Fukumura et al. (2010) found that pronoun use for a referent is reduced by the visual presence of another, competing referent which has not been mentioned in the linguistic context. Thus, Fukumura et al. (2010) argue that visual context information can become part of the discourse representation. Nevertheless, their study also shows that when both linguistic and visual information is presented, a referent's discourse status is more strongly affected by the linguistic context.

To summarise, the findings speak in favour of closely interconnected visual-perceptual and linguistic processing mechanisms. Primarily, accessibility of a referent's mental representation seems to depend on how it is transmitted linguistically (e.g., subject or topic status), whereas visual salience has been argued to have a weaker impact (Arnold & Lao, 2015; Bock et al., 2004; Fukumura et al., 2010) and/or might be considered at different processing levels (Vogels et al., 2013). However, the reviewed evidence showing that both topic status and visual cues of a social-communicative nature (e.g., speaker's gaze) guide sentence processing and pronoun resolution is based on highly diverse experimental paradigms and frameworks. Therefore it is still unclear, if and how a referent's accessibility in the mental model is modulated by inducing a referent as salient via linguistic or visual cues.

1.3. The present study

We aim to characterise the impact of a linguistic (Exp. 1: topic status) vs. visual salience cue (Exp. 2: gaze shift of a virtual person) on sentence processing and later pronoun resolution in German. Each of these cues was thought to increase the salience of a depicted referent in order to guide a referent's accessibility in discourse. In both experiments participants viewed a scene of potential referents (i.e., animals) performing a joint action. Afterwards the critical cue was presented using a between-subject design: In Experiment 1, the linguistic salience cue was realised by means of a context question that indicated one of the referents as the topic of the scene. In contrast, in Experiment 2, the visual salience cue was realised by a gaze shift of a virtual person to one of the depicted referents in order to draw the participant's attention to this referent. Given that people tend to look at what they are talking about and/or what they are currently attending to, it is valid to assume a functional similarity between linguistic topic marking and gaze – both indicate what the current communication is about. Therefore, we compared whether these functionally similar cues also affected sentence comprehension in a similar way. In either experiment, an equivalent neutral cue, indicating none of the animals as more salient amongst others, was presented in the respective modality.

To compare the impact of cue modality (linguistic vs. visual) on initial sentence processing and later pronoun resolution, the same experimental material and paradigm was used in both experiments. We assessed the effect of the salience cue on sentence processing by self-paced reading times during processing of sentences with varying word order. In addition, we assessed the effect of the salience cue on pronoun resolution by reading times during subsequent pronoun processing and antecedent choice preferences for this pronoun.

In the present study, we used the relatively flexible German word order that enables investigating the impact of salience across different sentence positions and independent of subject/agent status. Due to morphological case marking at the respective determiner phrase (DP), either the subject or object can be mentioned first in main clauses. The subject is marked with nominative case (NOM) while the object is marked with accusative case (ACC) (see example (1) vs. (2) for the SO and OS word order).

(1) Der grüne Fisch malt gleich den blauen Fisch vor den Blumen.
[the$_{[NOM]}$ green$_{[NOM]}$ fish$_{[NOM]}$]$_{subject/DP1}$ [paints]$_{verb}$ [now]$_{adverb}$ [the$_{[ACC]}$ blue$_{[ACC]}$ fish$_{[ACC]}$]$_{object/DP2}$ [in front of the flowers]$_{prepositional\ phrase}$.
"The green fish now paints the blue fish in front of the flowers."

(2) Den blauen Fisch malt gleich der grüne Fisch vor den Blumen.
[the$_{[ACC]}$ blue$_{[ACC]}$ fish$_{[ACC]}$]$_{object/DP1}$ [paints]$_{verb}$ [now]$_{adverb}$ [the$_{[NOM]}$ green$_{[NOM]}$ fish$_{[NOM]}$]$_{subject/DP2}$ [in front of the flowers]$_{prepositional\ phrase}$.
"The blue fish is now painted by the green fish in front of the flowers."

(3) Er freut sich schon auf das Picknick.
[He$_{[NOM]}$]$_{subject}$ [is looking forward]$_{verb}$ [already]$_{adverb}$ [to the picnic]$_{prepositional\ phrase}$.
"He is already looking forward to the picnic."

To investigate whether the effects found for initial sentence processing are still present during later pronoun resolution, the SO or OS sentence was followed by a pronoun sentence (see example (3)). The third person pronoun "*He*" could either be resolved to the subject or object of the preceding sentence.

We pursue the following line of reasoning based on the idea of mental models in which a referent's accessibility has been proposed to be modulated by both linguistic and visual-perceptual information (e.g., Johnson-Laird, 1980). If a linguistic or visual cue modulates the accessibility of a referent's mental representation, this should ease processing of sentences mentioning the salient referent first, as this position is typically attributed to highly accessible referents. Hence, for initial sentence processing we predict reduced reading times at sentence-initial position (i.e., DP1) for sentences mentioning the salient referent first compared to a preceding neutral cue that does not indicate any referent as more salient. Based on recent findings for linguistic contexts, reduced reading times for salient first-mentioned referents (i.e., topics) should be most visible in otherwise hard to understand OS sentences, as SO sentences are easy to process even in neutral contexts (Burmester et al., 2014). Based on previous evidence concerning the impact of visual-attentive cues on first-mention (e.g., Gleitman et al., 2007), visual salience might also ease processing of salient first-mentioned referents during subsequent sentence processing.

For later pronoun resolution, we predict that a linguistically or visually salient referent should be a good candidate for being chosen as the antecedent referent for the subsequent personal pronoun (i.e., "*He*"). We predict that the reliance on the default subject preference in pronoun interpretation might be reduced in favour of the object, if the pronoun is preceded by a salient-initial OS sentence, as the salience cue plus first-mention of the object might cause a competition of accessibility with the subject. If the impact of salience is still present during pronoun processing in a subsequent sentence, the impact of salience can be argued to be stable also at later processing stages. However, given previous evidence, the impact of the visual salience cue might not be as strong as expected for the linguistic salience cue. Still, if both cues (visual, linguistic) impact sentence processing and/or pronoun resolution, basal cognitive processes such as attention would need to be considered for future theories on information structure.

2. Materials and methods

In two experiments either a linguistic (Exp. 1: topic status) or visual cue (Exp. 2: gaze shift of a virtual person) indicated a visually depicted referent as more salient amongst others. Except for the cue modality, the material and the procedure were the same in both experiments.

2.1. Participants

Participants in Exp. 1 ($N = 27$, 14 female, *M* age 25.37 years, age range 19–34 years) and Exp. 2 ($N = 27$, 14 female, *M* age 25.01 years, age range 19–42 years) were German native speakers. None of them participated in both experiments. Three participants (i.e., two of Exp. 1 and one of Exp. 2) were excluded from analyses due to response accuracy scores below 60% in the probe comprehension questions or less than 50% of answered antecedent choice questions. The other 51 participants showed a mean response accuracy of 87% and responded to 98% of the antecedent choice questions, which indicates that these participants were attending to the experiment. Participants showed a normal reading span performance as measured by the German version of the standard computerised Reading Span Test of van den Noort, Bosch, Haverkort, and Hugdahl (2008) ($M = 63.1$, $SE = 1.54$, $CI = 3.03$). All but two ambidextrous participants were right-handed as assessed by a German version of the Edinburgh Handedness Inventory (Oldfield, 1971) and had normal or corrected-to-normal vision. None reported any neurological disorder. Participants received 11.50 € or course credits.

2.2. Design and material

In the two experiments, MODALITY of the cue (*linguistic* vs. *visual*) was included as a between-subject factor (Exp 1: *linguistic*, Exp 2: *visual*). In each experiment, a within-

subject design with the factors CUE (*salient vs. neutral,*) and WORD ORDER (*SO* vs. *OS*) was applied, respectively. For the *salient* CUE, the subsequently presented SO or OS sentence mentioned the salient referent sentence-initially (as DP1). To counterbalance salient-first sentences ($n = 60$) in which the salient referent was mentioned preverbally, the same amount of sentences but with the salient referent mentioned postverbally (as DP2) was used as filler sentences.

Previous studies addressing the impact of visual contexts on language processing showed that apart from salience, various additional factors are crucial, for instance, position of pictures in the visual display (e.g., tendency to initially fixate the left picture, Dahan et al., 2007; Gleitman et al., 2007) or visual depiction of theta roles of the sentence constituents (e.g., Knoeferle et al., 2005; Zhang & Knoeferle, 2012). Moreover, production studies, revealing an impact of visual cueing on first-mention, depicted referents in action such that theta roles were visible in the visual context (e.g., Gleitman et al., 2007; Myachykov et al., 2012). Therefore, we aimed to account for these two factors by controlling for ORIENTATION of animals in the visual scene (*left-to-right* vs. *right-to-left*) as well as for the predictability of the THETA ROLE of the first-mentioned referent (To improve readability, the results regarding these two control factors are reported in the Supplemental data). THETA ROLE was realised by displaying three animals of which the animals on the edge were either agent (oriented towards another animal) *or* patient (oriented away from another animal) (*theta role predictable*) and the central animal could be both, agent *and* patient of the action (*theta role unpredictable*) (see Figure 1). As animacy of sentence constituents also affects word order in German (i.e., tendency that animate precede inanimate referents; Grewe et al., 2006), we limited the linguistic material to animate referents that were all equally plausible to be the agent/subject or patient/object referent.

Each trial consisted of a short story depicting a triplet of contrastively coloured animals of the same type that are going to perform a joint action. To create 15 trials per condition, 15 action verbs (symmetric, transitive, 1-syllabic) were randomly combined with one of 15 animals (monomorphemic nouns of masculine gender, 1-syllabic ($n = 7$) and 2-syllabic ($n = 8$), *M* name agreement of 118 adults = 80%). All items were controlled for normalised lemma frequency values according to the dlex database (Heister et al., 2011). The animals were coloured in muted blue, green, red and/or yellow (1-syllabic colour adjectives in German) such that none of the animals of a triplet was more salient according to its colour.

In both experiments animals of a triplet were visually depicted in action one behind the other, all either looking from left-to-right or right-to-left (see Figure 1 (1)). As mentioned above, the animals on the edge were either agent (oriented towards another animal) *or* patient (oriented away from another animal) and the central animal could be both, agent *and* patient of the action.

Each trial contained a lead-in sentence (1), the cue context (2), the word order target sentence in SO or OS order (3), and the pronoun sentence (4) (see Figure 1). The lead-in (1) mentioned the animals of the story and drew participants' attention to all depicted animals. Thus, all animals of the scene were discourse-given in the linguistic as well as in the visual experimental paradigm, and hence their lemma representation was already activated before the cue context was presented. In Exp. 1 the lead-in showed the written text "*Look, there are three fish.*" (in the centre of the screen) and the animal triplet (at the lower part of the screen). During the presentation of the cue context (2) the animals remained the same but instead of the lead-in a wh-question was presented. The wh-question either induced a wide scope on the scene (*"What exactly is going on?",* neutral cue) or indicated one of the animals as the aboutness topic (*"What about the left/central/right one?",* salient cue). In Exp. 2 the lead-in (1) showed an animated female virtual person in frontal perspective blinking once with her eyes (presented on the upper part of the screen) plus the written text and animals as in Exp. 1. In the cue context (2) of Exp. 2, gaze and head movement of the virtual person were used instead of the wh-question to modify the salience of an animal. The virtual person either remained passive (neutral cue: no blinking, gaze or head movement) or turned her gaze and head towards one of the animals (salient cue).

The WORD ORDER of the target sentence (3) in either SO or OS order provided an answer to the preceding cue context by describing the thematic role relations of the acting animals (who was doing what to whom). The sentence consisted of a first determiner phrase (DP1) that was either subject and agent (marked with nominative case (NOM), see example (1), or object and patient (marked with accusative case (ACC), see example (2), followed by the verb, the second determiner phrase (DP2) that was either subject or object (inverse of DP1) and a closing prepositional phrase specifying the animals' location. The prepositional phrase was inserted to build a coherent story with the following pronoun sentence and prevent that reading times at DP2 are contaminated by "wrap up" effects due to sentence ending (e.g., Just & Carpenter, 1980).

The subsequent pronoun sentence (4) provided a continuation of the scene about one animal looking forward to some kind of action. The pronoun sentence

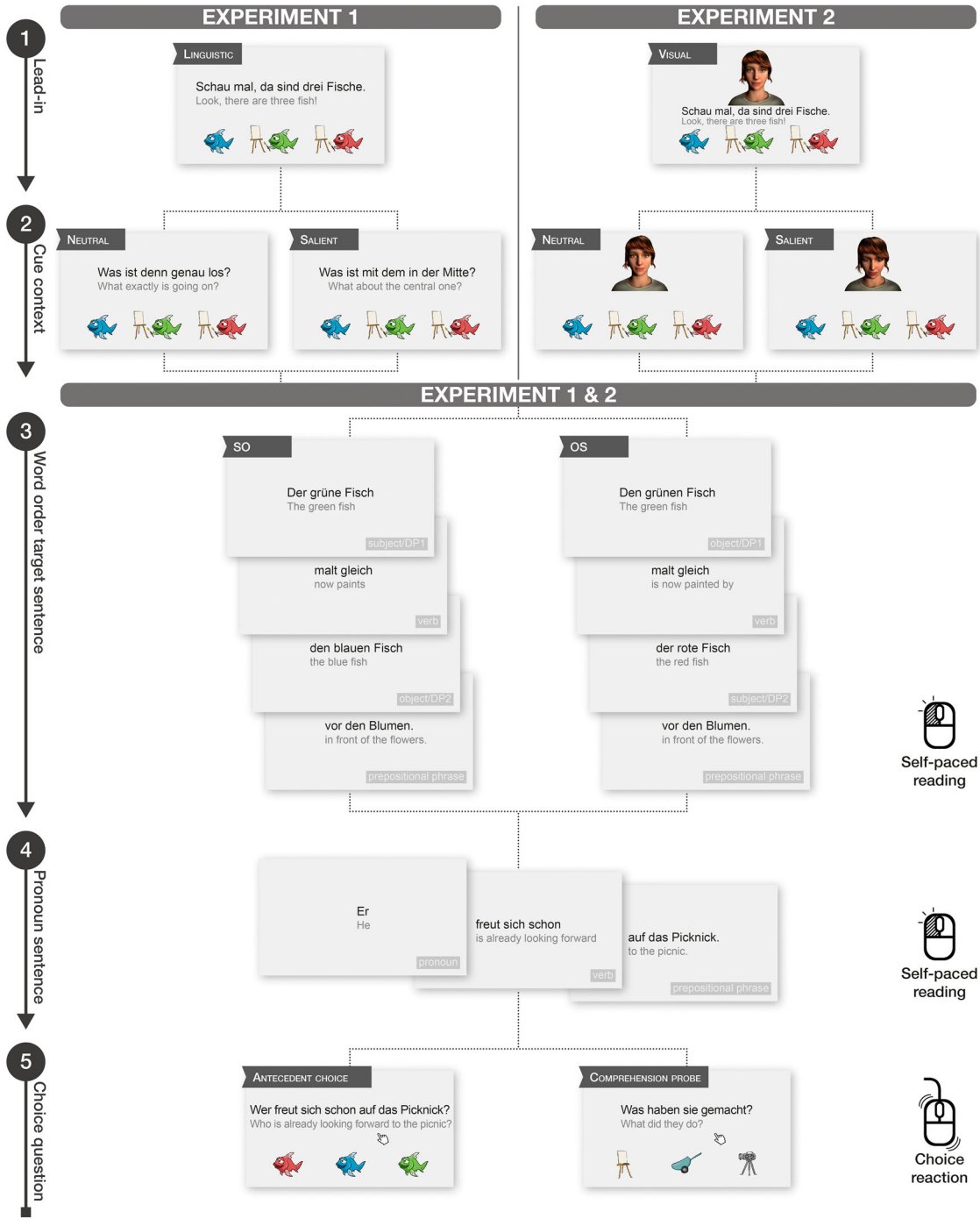

Figure 1. Experimental design of Exp. 1 (LINGUISTIC MODALITY) and 2 (VISUAL MODALITY) with approximate English translation written in grey. First, participants read the lead-in sentence (1). Subsequently, the salient or neutral CUE context (2) was presented followed by the SO or OS WORD ORDER target sentence (3). In the salient cue condition, the SO or OS sentence mentioned the salient referent (i.e., green fish) first (i.e., as DP1). Afterwards, the pronoun sentence (4) was presented. In the choice question phase (5), either an antecedent choice question or a comprehension probe had to be answered.

Abbreviations: SO = subject-verb-object, OS = object-verb-subject, DP = determiner phrase.

always started with the masculine third person pronoun "*He*" (marked with nominative case (NOM), see example (3) for which both subject and object of the preceding sentence could be plausible antecedent referents. Each story was randomly followed by an antecedent choice question in ten trials per condition *(e.g., "Who is looking forward to the picnic?")* or a comprehension probe in five trials per condition which asked for the agent *("Who was just painting?")*, patient *("Who was just being painted?")*, location *("Where are they?")*, excitement *("What is he looking forward to?")*, or action *("What did they do?")* (see Figure 1 (5)).

In total each participant read 180 different stories. Within the experiments each participant was presented with a different pseudo-randomised order (criteria: maximally two consecutive trials of the same condition and word order, maximally four consecutive trials with the salient animal in the same location (left, central, right) and an equal likelihood of the four colours representing the salient animal). Within each condition the animals equally often had a left-to-right vs. right-to-left orientation.

2.3. Procedure

Participants were tested individually, seated in front of a computer screen on which the experiment was displayed by means of the Presentation® software (version 18.1, www.neurobs.com). The participant was introduced to the experimental procedure by a written instruction on the screen presented in a speech bubble next to the virtual person depicted in Exp. 2. The virtual person was programmed by using the DAZ 3D Studio® software (version 4.6, www.daz3d.com). To embed the experiment in a cover story and make the stories about animals plausible, the virtual person informed the participants that she usually tells these stories to children. Prior to each trial a red cross in the centre of the screen signalled the beginning of a new story. Via mouse click (with the right index finger) lead-in and cue context were presented in a preset time window of three seconds each. A black fixation cross in the centre of the screen signalled the start of the self-paced reading sequence. Via consecutive mouse clicks, the word order target and pronoun sentence were presented phrase-wise (see Figure 1 (3) and (4)). Participants were instructed to read and look at each story attentively and silently, to read the phrase-wise presented sentences as naturally as possible and to answer the choice question after each story (i.e., antecedent choice question or comprehension probe) as accurately and fast as possible within a four second time window (see Figure 1 (5)). Participants were made aware that for some questions there is no correct or incorrect answer, but that they should judge on their intuition in these cases. Participants responded to the choice question via mouse click on one of the three pictures (e.g., animals depicted without instruments and in randomised order, depicted action). To become familiar with the procedure participants performed five practice trials. Reading times for the phrase-wise presented sentence positions, responses of the choice questions (i.e., antecedent choices and response accuracy for comprehension probes), and their response times were recorded. The experimental session lasted about 45 min including two pauses and was followed by a short questionnaire concerning strategic behaviour.

2.4. Data analysis

For the statistical data analysis, linear mixed effects models were calculated using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) provided by the *R* environment (version 3.2.3., R Core Team, 2015). Linear mixed effects models were calculated to assess the impact of Modality, Cue, Word order, and the control factors Theta role and Orientation as fixed effects and Participants and Items as random effects. The two-level factors Modality *(linguistic* vs. *visual)*, Cue (*salient* vs. *neutral*), and Word order (*SO* vs. *OS* sentences) as well as the control factors Theta role (*predictable* vs. *unpredictable*) and Orientation *(left-to-right vs. right-to-left)* were coded as +/−.5 to resemble the contrast (sum) coding of traditional ANOVA analyses. As estimating maximal fitted models might not be sufficient for our data and led to overparameterisation as indicated by convergence errors (Bates, Kliegl, Vasishth, & Baayen, 2015), model fitting started with the simple model; that is with all fixed effects and their interactions, and Participants and Items defined as random intercepts. In a step-wise

Table 1. Linear mixed effects model output for log reading times across sentence positions of the word order sentence.

	DP1			verb			DP2		
FIXED EFFECTS	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
intercept	6.668	0.034	**196.51***	6.426	0.037	**173.90***	6.610	0.045	**147.04***
Modality	−0.108	0.068	−1.59	−0.033	0.073	−0.45	−0.184	0.089	**−2.05***
Cue	0.066	0.013	**5.09***	0.013	0.008	1.60	0.034	0.010	**3.51***
Word order	−0.060	0.010	**−5.93***	−0.036	0.010	**−3.49***	−0.112	0.019	**−6.02***
Modality × Cue	0.042	0.026	1.60	−0.008	0.016	−0.48	−0.021	0.019	−1.10
Modality × Word order	0.012	0.020	0.58	0.018	0.020	0.90	0.074	0.037	**1.98***
Cue × Word order	−0.018	0.015	−1.19	−0.005	0.016	−0.34	−0.011	0.019	−0.56
Modality × Cue × Word order	−0.011	0.030	−0.35	−0.014	0.031	−0.44	−0.005	0.038	−0.13
Formula of final model with random slope adjustments for Participants (P) and Items (I)	DP1 ~ Modality * Cue * Word order * Theta role * Orientation * + (1+ Cue + Word order + Theta role | P) + (1 | I)			verb ~ Modality * Cue * Word order * Theta role * Orientation * + (1+ Word order | P) + (1+ Theta role | I)			DP2 ~ Modality * Cue * Word order * Theta role * Orientation * + (1+ Word order | P) + (1+ Theta role | I)		

Abbreviations: * = statistically significant effects with $|t| > 1.96$, *b* = estimate, *SE* = standard error, DP = determiner phrase.

Table 2. Linear mixed effects model output for log reading times across sentence positions of the pronoun sentence.

	pronoun			verb			prepositional phrase		
FIXED EFFECTS	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
intercept	6.317	0.040	**159.30***	6.306	0.040	**157.98***	6.458	0.036	**180.74***
MODALITY	0.101	0.079	1.28	−0.098	0.080	1.23	0.008	0.071	0.11
CUE	−0.005	0.013	−0.35	0.010	0.008	1.23	0.0003	0.009	0.03
WORD ORDER	−0.016	0.010	−1.66	−0.029	0.008	**−3.54***	−0.027	0.009	**−3.06***
MODALITY × CUE	−0.017	0.016	−1.08	0.002	0.016	0.12	0.007	0.018	0.42
MODALITY × WORD ORDER	−0.041	0.016	**−2.56***	−0.009	0.016	−0.57	−0.012	0.018	−0.69
CUE × WORD ORDER	−0.015	0.016	−0.93	−0.009	0.016	−0.53	−0.011	0.018	0.64
MODALITY × CUE × WORD ORDER	−0.043	0.032	−1.32	−0.013	0.033	−0.38	−0.053	0.036	−0.48
Formula of final model with random slope adjustments for Participants (P) and Items (I)	pronoun ~ MODALITY * CUE * WORD ORDER * THETA ROLE * ORIENTATION * + (1 | P) + (1+ CUE + WORD ORDER * THETA ROLE | I)			verb ~ MODALITY * CUE * WORD ORDER * THETA ROLE * ORIENTATION * + (1 | P) + (1 | I)			prepositional_phrase ~ MODALITY * CUE * WORD ORDER * THETA ROLE * ORIENTATION * + (1 | P) + (1 | I)		

Abbreviations: * = statistically significant effects with $|t| > 1.96$, *b* = estimate, *SE* = standard error.

manner, slope-adjustments were included if they significantly improved the explanatory power of the simpler model without that slope adjustment (Baayen, 2008). The final model with the final random effect structure for each dependent variable is reported in Tables 1–3.

The dependent variables for sentence processing were reading times of DP1, verb, and DP2 of the word order target sentence. The dependent variables for pronoun resolution were reading times of the pronoun and the spillover region (the subsequent verb and prepositional phrase) in the pronoun sentence as well as subject antecedent choices in the antecedent choice task. For statistical analyses of reading times per sentence position, the logarithmic (log) transformation was most suitable as determined by the *boxcox* function of the *MASS* package in *R* (Box & Cox, 1964; Venables & Ripley, 2002). For statistical analyses of antecedent choices, participants' responses were treated as binomial such that subject antecedent choices were coded as 1 (subject antecedent) or 0 (non-subject antecedent: object or other depicted referent). Concentrating on subject antecedent choices and deviations of that preference was motivated by the fact that in our study antecedent choices were highest for subjects ($M = 0.79$) which conforms to the generally reported subject antecedent preference (e.g., Bouma & Hopp, 2007 for German). For analyses of subject antecedent choices logit mixed models were computed (cf. Jaeger, 2008). For the final models, the statistics of the fixed effects are reported with estimates *(b)*, standard errors *(SE)*, and *z*- and *p*-values for binomial data, or *t*-values for reading time data. Effects are considered as significant at $\alpha = .05$ if $|z|/|t| > 1.96$. Significant interaction effects were resolved by calculating post hoc pairwise comparisons on the final mixed effects model (with the *R* package *lsmeans* (cf. Lenth, 2016)) for which we report *b, SE, z*- or *t*-values, and Tukey adjusted *p*-values.

Table 3. Logit mixed effects model output for subject antecedent choices.

	subject antecedent choices		
FIXED EFFECTS	*b*	*SE*	*z*
intercept	2.173	0.237	**9.15*****
MODALITY	−0.709	0.470	−1.51
CUE	0.249	0.106	**2.35***
WORD ORDER	1.391	0.329	**4.22*****
MODALITY × CUE	0.078	0.212	0.37
MODALITY × WORD ORDER	1.649	0.643	**2.56***
CUE × WORD ORDER	−0.629	0.212	**−2.97****
MODALITY × CUE × WORD ORDER	−1.278	0.424	**−3.02****
Formula of final model with random slope adjustments for Participants (P) and Items (I)	subject antecedent ~ MODALITY * CUE * WORD ORDER * THETA ROLE * ORIENTATION *+ (1+ WORD ORDER | P) + (1 | I)		

Abbreviations: Significance levels: *$p < .05$, **$p < .01$, ***$p < .001$, *b* = estimate, *SE* = standard error.

3. Results

Figure 2 illustrates the log reading times over the time course of relevant positions during processing the word order sentence and the pronoun sentence. With regard to pronoun resolution, Figure 3 shows the mean proportion of subject antecedent choices across conditions. The statistics of the fixed effects as revealed by the linear mixed effects models are reported in Tables 1–3. For post hoc models the significant effects are reported in the text with relevant values in brackets. Statistically significant effects involving the control factors THETA ROLE of the referent (predictable vs. unpredictable) as well as ORIENTATION of the depicted animals in the visual scene (looking from left-to-right vs. right-to-left) are reported in the Supplemental data.

3.1. Sentence processing

Statistical results as revealed by the linear mixed effects models of log reading times across sentence positions in the word order target sentence are provided in Table 1.

DP1. Crucially for our research question, MODALITY of cue did *not* significantly affect reading times of the

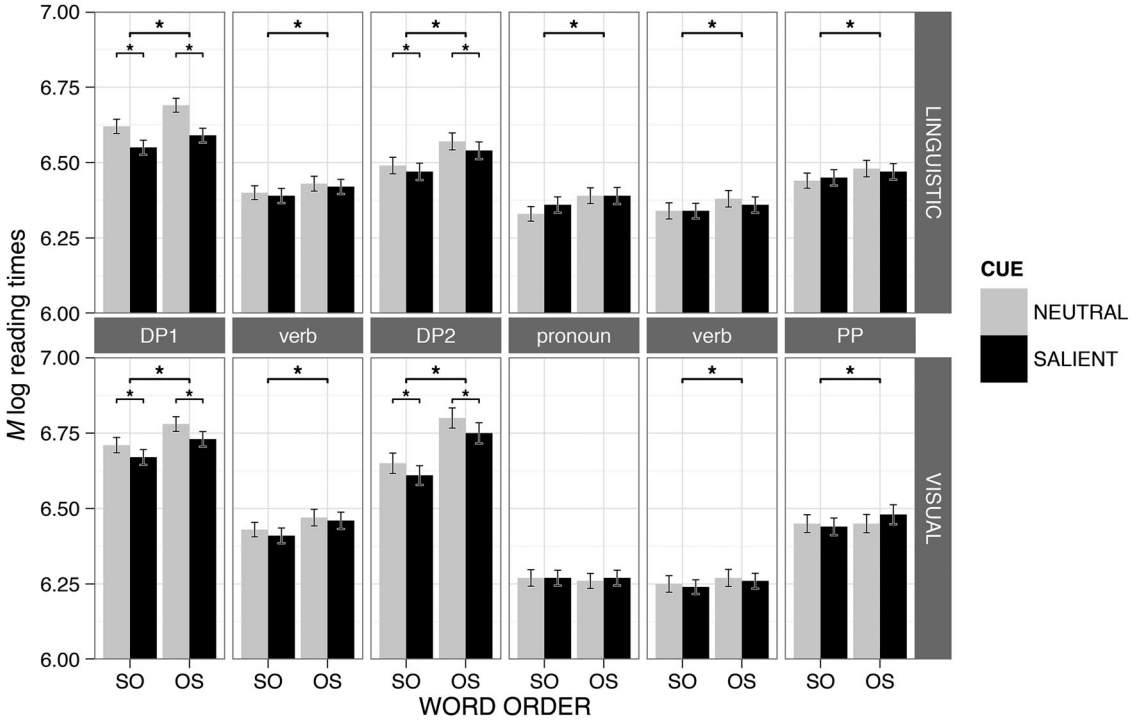

Figure 2. Mean *(M)* log reading times across relevant sentence positions during processing the word order sentence (DP1, verb, DP2) and the pronoun sentence (pronoun, verb, PP) for the LINGUISTIC (Exp. 1: upper panel) and VISUAL MODALITY (Exp. 2: lower panel). Statistically significant effects with $|t| > 1.96$ are marked with an asterisk. Error bars indicate 95% confidence intervals (*CI*s). For the within-subject factors (CUE, WORD ORDER) the *CI*s exclude the between-participant variance (Cousineau, 2005) and were corrected according to Morey (2008).

Note: Significant effects of MODALITY are not indicated in the figure but are only discussed in the main text (see section 3.1). Abbreviations: SO = subject-verb-object, OS = object-verb-subject, DP = determiner phrase, PP = prepositional phrase.

first-mentioned referent (DP1) directly following the linguistic or visual cue. Instead, log reading time analyses of DP1 revealed a significant main effect of CUE (salient vs. neutral) such that, independent of cue MODALITY, DP1 was read faster if it was previously indicated as salient in the linguistic and in the visual condition (see Figure 2). In addition, a significant main effect of WORD ORDER was reflected in faster reading times of DP1 in SO compared to OS sentences.

Verb. Statistical analyses of log reading times at the verb position revealed no modulation by the preceding linguistic or visual salience cue. Instead, analyses yielded a significant main effect of WORD ORDER such that reading times for the verb were significantly faster in SO compared to OS sentences.

DP2. Postverbally, at position of DP2, log reading time analyses revealed a significant main effect of MODALITY such that DP2 was read faster if preceded by a linguistic compared to a visual cue. Further, the statistical analysis showed a significant main effect of CUE such that DP2 was read faster if the preceding sentence-initial referent was previously indicated as salient compared to the neutral condition. Similar to DP1 and the verb, a significant main effect of WORD ORDER was reflected in faster reading times of DP2 in SO compared to OS sentences. However, the significant main effect of WORD ORDER was modulated by the significant interaction of MODALITY × WORD ORDER. Post hoc comparisons showed that the effect of WORD ORDER was present in both modalities, but with a lower impact in the LINGUISTIC MODALITY ($b = -0.076$, $SE = 0.027$, $t = -2.83$, $p = .033$) compared to the VISUAL MODALITY ($b = -0.150$, $SE = 0.026$, $t = -5.71$, $p < .0001$) (see Figure 2). Hence, across sentence positions WORD ORDER significantly affected sentence processing.

3.2. Pronoun resolution

3.2.1. Reading times of the pronoun sentence

Log reading time analyses for the sentence-initial pronoun "*He*" directly following the word order target sentence revealed a significant interaction of MODALITY × WORD ORDER (see Figure 2 for the plotted log reading times and Table 2 for the statistics of the fixed effects of the linear mixed effects models). As confirmed by post hoc comparisons, the WORD ORDER effect was significant for the LINGUISTIC ($b = -0.037$, $SE = 0.012$, $t = -3.20$, $p = .008$) but not for the VISUAL MODALITY ($b = 0.004$,

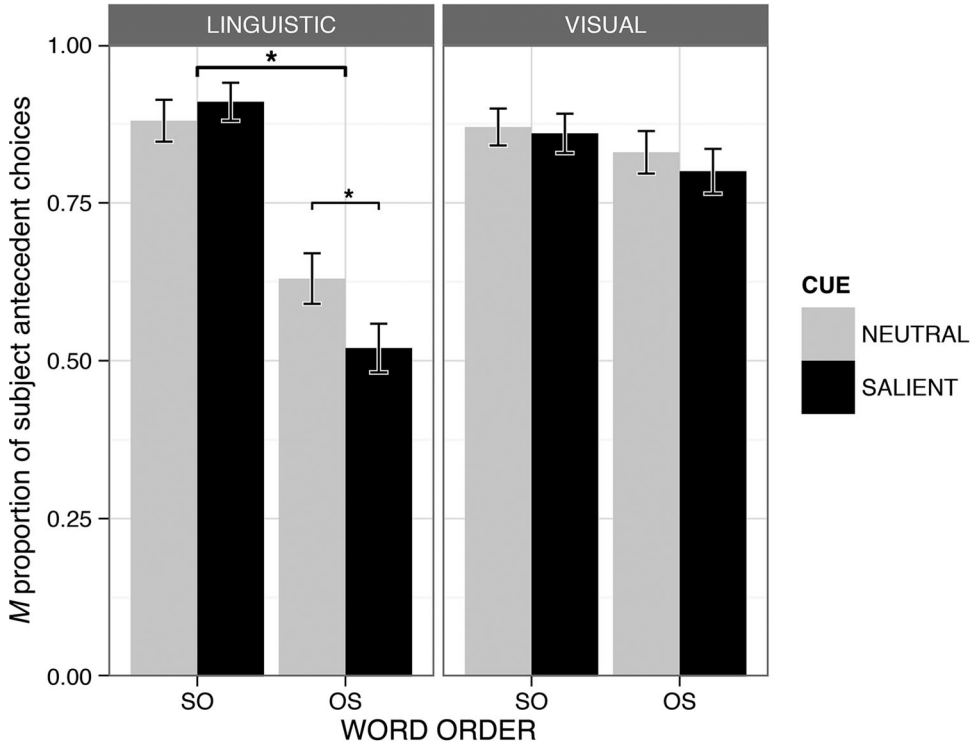

Figure 3. Mean *(M)* proportion of subject antecedent choices for the pronoun "*He*" for the LINGUISTIC (Exp. 1: left panel) and VISUAL MODALITY (Exp. 2: right panel). Statistically significant effects with $|z| > 1.96$ are marked with an asterisk. Error bars indicate 95% confidence intervals (*CI*s). For the within-subject factors (CUE, WORD ORDER) the *CI*s exclude the between-participant variance (Cousineau, 2005) and were corrected according to Morey (2008).

Abbreviations: SO = subject-verb-object, OS = object-verb-subject.

SE = 0.011, *t* = 0.38, *p* = .982). Thus, following the linguistic cue modality, the pronoun was read faster if preceded by SO compared to OS sentences; whereas following the visual cue modality, reading times for the pronoun were not modulated by the preceding word order.

As usual for reading time studies, we did not just look at the region of interest (i.e., the pronoun) but also investigated directly following sentence positions to catch possible spillover effects (c.f., Mitchell, 1984). With regard to the spillover region of the pronoun, reading times of neither the subsequent verb nor prepositional phrase were affected by cue MODALITY. However, the analyses revealed a significant main effect of WORD ORDER for the verb following the pronoun as well as for the sentence-final prepositional phrase. Hence, the verb and the prepositional phrase were read faster if preceded by an SO sentence compared to an OS sentence.

3.2.2. Antecedent choice preferences

The descriptive statistics show that in both experiments (i.e., both in the linguistic and in the visual modality), the referent in subject position was the preferred antecedent for the subsequent third person pronoun "*He*" (Exp. 1 (LINGUISTIC): *M* = 0.735, *SE* = 0.010; Exp. 2 (VISUAL): *M* = 0.841, *SE* = 0.008). The object referent was less likely to be chosen as the antecedent (Exp. 1 (LINGUISTIC): *M* = 0.251, *SE* = 0.010; Exp. 2 (VISUAL): *M* = 0.149, *SE* = 0.008). The other not mentioned animal was the least likely antecedent (Exp. 1 (LINGUISTIC): *M* = 0.013, *SE* = 0.003; Exp. 2 (VISUAL): *M* = 0.010, *SE* = 0.002). Note that grammatical function and theta role always coincided in our study (i.e., the grammatical subject was always the agent and the grammatical object was always the patient of the action).

Figure 3 shows the mean proportion of subject antecedent choices for the pronoun across conditions. Statistical analyses of subject antecedent choices yielded a significant main effect of CUE and of WORD ORDER such that subject antecedent choices were more likely following salient-first compared to neutral cues and following SO compared to OS sentences. However, these main effects were modulated by significant interactions of CUE × WORD ORDER, of MODALITY × WORD ORDER, and of MODALITY × CUE × WORD ORDER (see Table 3 for the logit mixed model output). Hence, depending on the cue modality, the likelihood of subject antecedent choices differed with regard to the impact of the salience cue and word order of the preceding sentence.

Post hoc comparisons showed that the linguistic salience cue affected subject antecedent choices only if presented prior to a non-canonical OS sentence (*b* =

0.922, $SE = 0.186$, $z = 4.95$, $p < .0001$), but *not* if the linguistic salience cue was presented prior to a canonical SO sentence ($b = -0.346$, $SE = 0.235$, $z = -1.47$, $p = .822$). As the mean proportion of subject antecedent choices plotted in Figure 3 shows, the significant effect of the linguistic salience cue was reflected in lower subject antecedent choices following salient-first OS sentences compared to OS sentences in the neutral condition. In contrast, the visual salience cue did not significantly affect subject antecedent choices, neither if presented prior to SO sentences ($b = 0.215$, $SE = 0.227$, $z = 0.95$, $p = .981$) nor if presented prior to OS sentences ($b = 0.205$, $SE = 0.195$, $z = 1.05$, $p = .967$). In addition, the post hoc comparisons showed that with regard to the linguistic cue subject antecedent choices were more likely following SO compared to OS sentences, both when the linguistic cue indicated one referent as salient ($b = 2.849$, $SE = 0.487$, $z = 5.85$, $p < .0001$) and when the linguistic cue was neutral ($b = 1.581$, $SE = 0.482$, $z = 3.28$, $p = .023$). With regard to the visual cue, subject antecedent choices were not modulated by the preceding word order, neither in the salient cue condition ($b = 0.561$, $SE = 0.480$, $z = 1.17$, $p = .941$) nor in the neutral cue condition ($b = 0.572$, $SE = 0.488$, $z = 1.17$, $p = .940$).

To summarise, following the linguistic salience cue, subject antecedent choices were modulated by the preceding salience cue and word order. In contrast, following the visual salience cue, subject antecedent choices were not modulated by the preceding cue and/or word order.

4. Discussion

We aimed to compare effects of cue modality on the degree of a referent's mental accessibility by means of two experiments in which we evaluated the impact of linguistic vs. visual salience on sentence processing and later pronoun resolution in German. As salient referents are very likely to be mentioned in sentence-initial position, we expected to see the strongest effect of the cue at the sentence-initial position (i.e., DP1). Hence, with regard to sentence processing we investigated reading times during processing of SO and OS sentences mentioning the salient referent first (i.e., as the subject or object). Moreover, as salient referents have a strong tendency to be interpreted as the antecedent referent during pronoun resolution, we investigated reading times and antecedent choice preferences of a subsequent personal pronoun. Our findings indicate that for sentence processing, both the linguistic and visual salience cue immediately speeded reading times of the salient first-mentioned referent in the subsequent SO and OS sentence. For pronoun resolution, the results show a less congruent effect of both cue modalities. Furthermore, across all conditions (i.e., in both cue modalities and during sentence and pronoun processing), significant interactions with the control factors theta role and orientation (i.e., position of the animals in the visual scene) have shown to be relevant during sentence comprehension (see Supplemental data for details). In the following, our findings will be discussed with a focus on the impact of cue modality on sentence processing and pronoun resolution.

4.1. The impact of linguistic vs. visual salience on sentence processing

With regard to sentence processing, our results demonstrate that if a sentence-initial referent has previously been indicated as more salient amongst others –by a linguistic or visual cue– processing of the subsequent sentence profited from the salience cue. Independent of the canonical or non-canonical word order the salience cue eased sentence processing. That is, if the most salient and hence most accessible referent was first-mentioned, reading times at sentence-initial position (i.e., DP1) directly following the salience cue were faster compared to a preceding neutral cue. This processing advantage due to the salience cue was still present at the position of the last-mentioned referent, as reflected in faster reading times of DP2. However, at DP2 the linguistic cue speeded reading times more than the visual cue. Crucially, the results of the present study substantiate previous evidence on the impact of linguistic context on German sentences with varying word order (e.g., Bornkessel & Schlesewsky, 2006; Burmester et al., 2014; Weskott et al., 2011). While Burmester et al. (2014) showed that the processing of non-canonical sentences is enhanced if the sentence-initial referent is indicated as more salient by a directly preceding linguistic (aboutness topic) context, our current study expands these findings to the impact of linguistic and visual salience on the processing of both, canonical as well as non-canonical sentences.

As expected, canonical sentences were easier to process compared to non-canonical sentences in terms of faster reading times across all positions of the word order target sentence supporting the well-established word order effect in German (e.g., Matzke, Mai, Nager, Rüsseler, & Münte, 2002). Moreover, if the first-mentioned referent was linguistically or visually cued, readers profited from the predictability of theta role information based on the depicted animals in action (see Supplemental data for the results). In sum, during sentence processing we see similar facilitating effects of linguistic and visual cues.

4.2. The impact of linguistic vs. visual salience on pronoun resolution

As predicted based on multiple studies on pronoun resolution (e.g., Bosch et al. 2007; Bouma & Hopp, 2007), the personal pronoun in our study was preferentially interpreted in favour of the subject of the preceding sentence following both, the linguistic and visual cue. However, our findings show that linguistic and visual cues modulated antecedent choices differently: In the visual modality, neither the preceding salience cue nor word order influenced participants in their preference to choose the subject as antecedent referent for the pronoun. In contrast, in the linguistic modality, the subject antecedent preference was modulated by the salience cue and the word order of the preceding sentence: Subject antecedent choices were reduced following OS compared to SO sentences. In addition, subject antecedent choices were reduced if preceded by salient-first OS sentences compared to OS sentences presented in neutral contexts. That is, the strong preference to interpret pronouns in favour of the highly accessible subject (and agent) referent of the preceding sentence was modulated by linguistically induced salience of the object. Hence, linguistic salience (due to topic status and/or first-mention) had the potential to increase the accessibility of referents with a less prominent grammatical role (i.e., objects) – albeit to a lesser degree than subject (and agent) status. These findings are in line with previous research supporting the crucial role of the object referent's topic status in the processing of OS sentences (Burmester et al., 2014) and with regard to pronoun resolution in German (Colonna et al., 2012). In line with Colonna et al. (2012), German personal pronouns were resolved in favour of the preceding subject by default, that is, independent of additional topic features or first-mention of the subject. By contrast, in the non-canonical word order, topicalised objects increased referent accessibility as reflected by increased object antecedent interpretations compared to the canonical word order, although to a significantly lesser degree than topicalised subjects (Colonna et al., 2012). Moreover, the impact of topic status on pronoun resolution in our study supports the findings of Bosch and Umbach's (2007) corpus and reading time study with regard to the important role of information-structural properties (i.e., topic status) of the potential antecedent referent.

With regard to reading times of the pronoun sentence, the cue modality also seemed to cause a different processing at position of the pronoun: Reading times of the pronoun were only modulated following the linguistic but not the visual cue: Following the linguistic cue, both when the cue was neutral or salient, reading times were faster following SO compared to OS sentences. However, during subsequent processing (i.e., of the spillover region), the word order effect was present in both modalities.

In sum, following the linguistic cue, both subject antecedent choices and pronoun reading times were influenced by the preceding salience cue and/or word order. In contrast, the visual cue did not produce a comparable impact on pronoun resolution, as neither subject antecedent choices nor pronoun reading times were influenced by the visual salience cue and/or word order. However, as discussed in the next section, pronoun resolution might not reflect the pure impact of linguistic vs. visual salience: In our study additional linguistic material that preceded the pronoun (i.e., the word order target sentence; e.g., "*The green fish now paints the blue fish in front of the flowers.*") also affected pronoun resolution.

4.3. General discussion

Our findings concerning sentence processing and pronoun resolution make an important contribution to the understanding of discourse in terms of a situational setting in which the interplay of linguistic as well as visual information is decisively considered in the listener's mental representation. The comparable impact of linguistic and visual salience cues on reading times (i.e., sentence processing) vs. their unequal impact on antecedent choice preferences (i.e., pronoun resolution) suggest that visual cues contribute to a referentś accessibility differently. It looks as if the impact of visual salience in our experimental design was restricted to the initial reading times – that is, the visual salience cue facilitated immediate sentence processing, maybe by modulating hearers' expectations about the next mentioned element. The impact of linguistic salience, however, seems to have a longer lasting impact as it affected not only initial processing but also subsequent interpretative processes in a given trial: Pronoun resolution was only affected by the linguistic –but not the visual– cue (in combination with word order) to counteract the bias to choose the subject as antecedent.

Various accounts have dealt with the differential impact of linguistic and visual cues. The production study by Vogels et al. (2013) revealed that a referent's visual salience makes this referent more likely to be mentioned first but does not impact if a pronoun is used to refer to it. In contrast, in their study, a referent's linguistic salience affects both, first mention and pronoun use. They argue that in contrast to linguistic salience, visual salience only affects the global interpretation of the

scene in the sense of identifying which referent is most important and is thus mentioned first. Still, according to Vogels et al., visual salience does not affect the accessibility of a referent's mental representation as would have been reflected in frequent pronoun use. Opposed to Vogel et al.'s interpretation, others argue that already the visual presence of other depicted referents causes a competition of referent accessibility in discourse which lead to a reduced pronoun use (Arnold & Griffin, 2007; Fukumura et al., 2010).

However, if both visual and linguistic information is presented simultaneously (as in our study), linguistic salience has found to more strongly affect referent accessibility than a referent's visual salience: Fukumura et al. (2010) suggested that the more salient a referent is in the linguistic context (i.e., due to subject status), the weaker is the impact of its visual salience. This line of argumentation is supported by research on the comprehension of ambiguous personal pronouns similar to our study (e.g., Arnold & Lao, 2015; Nappa & Arnold, 2014): Therein visual salience cues have shown to be only of secondary relevance for pronoun resolution whereas linguistic information (i.e., first-mention in the preceding sentence) played the primary role. Moreover, as evidenced previously, visual information might only be considered if linguistic information is uninformative or ambiguous in the situational context (Nappa, Wessel, McEldoon, Gleitman, & Trueswell, 2009). Note that in our study the word order target sentence was directly preceded by the visual cue. Therefore the visual salience cue might have induced a similar effect during sentence processing as the linguistic salience cue. In contrast, the pronoun was directly preceded by the word order sentence so that -in case of the visual cue- linguistic information might have interfered. We follow the line of argumentation of previous research that linguistic salience weakens the impact of visual salience (e.g., Fukumura et al., 2010). In our study, linguistic salience in terms of grammatical subject and agent status together with being first-mentioned in the directly preceding sentence yielded the stronger impact relative to the visual salience cue when it came to a competition of referent accessibility. Alternatively, according to Nappa et al. (2009), the word order sentence in our study was already informative enough to identify the antecedent referent such that the visual cue was less strongly considered for pronoun resolution. This issue might be solved by a paradigm in which linguistic or visual salience cues are directly followed by pronominal coreference as for instance exemplified by Ariel (1988): A scene in which someone suddenly leaves a meeting followed by the utterance "*He must be upset.*" (p. 80, footnote 12).

Notably, gaze is a powerful but ambiguous cue, because it is not always related to the message content or tightly coincided with speech (Hanna & Brennan, 2007). Likewise gaze corresponds to a variety of mental states that might be difficult to capture (Staudte & Crocker, 2011). Nevertheless, gaze has been shown to help listeners to identify the speaker's focus of attention as well as to understand speaker's intentions by for instance anticipating what the speaker is going to talk about next, and to finally facilitate comprehension mechanisms (Knoeferle & Kreysa, 2012). Note that our findings are based on written sentence comprehension which might narrow down the impact of the visual cue compared to spoken comprehension which would be closer to natural communicative settings of interlocutors. Hence, visual cues such as gaze might be easier accessible during spoken than written language processing. Moreover, previous research points out that visual cues (e.g., gestures, eye gaze) only affect sentence processing and pronoun resolution if these cues serve a clear communicative function in the listeners mind; in case of more abstract visual cues (e.g., moving point, unconscious flash), speakers' intentions are not inferable and hence comprehension processes would not profit (Holle et al., 2012; Nappa & Arnold, 2014). Our design cannot distinguish whether the visual salience cue eased processing of the salient first-mentioned referent due to a modulation of participants' visual attention to the depicted referent or due to the intentional character of the gaze cue. However, the visual cue similar to the linguistic cue immediately increased referent accessibility during sentence processing such that participants profited from the salience cues in terms of processing speed.

Future research is needed to disentangle some of the above mentioned alternative explanations such as the influence of the type of the (especially visual) cue context, the impact of the modality of the presented sentences (spoken vs. written), and the actual processing level or possibly hierarchy with which the different visual and linguistic features contribute to the accessibility of a referent's mental representation. Moreover, the modulation of visual salience in the form of a virtual person's gaze might be a less compelling social cue as of a real person in face-to-face communication in which speaker's gaze is usually aligned with speech. Furthermore, other populations such as individuals with language disorders or children might be more sensitive to visual (gaze) cues than healthy young adults as for them non-linguistic cues might be more important for language comprehension processes and social interaction.

In sum, our findings substantiate a close coupling of linguistic and visual information supporting the

assumptions of the *Coordinated Interplay Account* (Crocker et al., 2010) in terms of closely interconnected processing mechanisms of both. According to the Coordinated Interplay Account, linguistic information guides listeners' (visual) attention to information in the visual environment (cf. utterance-mediated attention) and vice versa, visual, scene-based information affects sentence processing in a temporally close relationship. This view of situated sentence processing considering context-specific information of the linguistic as well as visual domain is supported by a bunch of recent research investigating the comprehension and production of sentences and pronouns while recording listener's (or speaker's) eye movements or electrophysiological responses (e.g., Arnold & Lao, 2015; Knoeferle et al., 2007). As discussed along with the framework of the Coordinated Interplay Account, many theories on sentence processing do not explicitly consider the role of extra-linguistic, visual, information, which is also true for theories on referential processing (see Crocker et al., 2010 for the discussion of syntactic processing models therein). Moreover, the assumption that language comprehension comprises of more than purely language-mediated processing mechanisms is shared by the *Syntax-Discourse Model* (Schumacher, 2014). As supported by electrophysiological responses, the Syntax-Discourse Model emphasises the role of context information during the two mechanisms of meaning computation. During the first mechanism (so-called "discourse-linking") the listener builds a discourse representation based on prior context which is understood as a broad notion of (amongst others) sentential context, situational context, and world knowledge. The second mechanism ("discourse updating") draws inferences in order to come up with a feasible interpretation based on a cooperative speaker-hearer interaction. Hence, the model assumes that language comprehension relies on context-dependent pragmatic information while stressing the crucial role of the hearer's assumptions about the speaker's intention. The finding of our study that linguistic and visual salience cues modulated sentence processing and pronoun resolution –although in a different way– supports the assumption that multimodal information is reconciled in the readers' mental representation of discourse.

5. Conclusion

Our study aimed to compare the impact of a linguistic vs. visual salience cue on sentence processing and later pronoun resolution. In sum, both the linguistic and visual salience cues immediately speeded up reading times of German canonical and non-canonical sentences mentioning the salient referent first. Hence, for sentence processing we can conclude that readers were similarly sensitive to both the linguistic and visual cue indicating a depicted referent as more salient amongst others. Concerning pronoun resolution, linguistic cues affected antecedent choices differently to visual cues: Following the linguistic salience cue, the strong subject antecedent preference was reduced following non-canonical compared to canonical sentences, whereas linguistic salience (i.e., topic status) of the object referent increased its likelihood of being interpreted as the antecedent of the following pronoun. In contrast, following the visual cue, readers relied on their default strategy to interpret pronouns in favour of the subject of the preceding sentence.

Taken together, our findings provide further evidence that a referent's mental accessibility is influenced by its linguistic and visual salience in discourse. Reader's sensitivity to linguistic vs. visual salience cues differed with regard to sentence processing and pronoun resolution as for the latter linguistic features played the primary role. Finally, our findings are in line with accounts of language comprehension that point to the impact of linguistic and visual information (amongst others) and share the assumption of closely interconnected processing mechanisms of context information delivered via different input modalities (Crocker et al., 2010; Schumacher, 2014).

Notes

1. Salience has also been described in earlier frameworks, e.g., Praguès school (e.g., Sgall, Hajičová, & Panevová, 1986), *Centering Theory* (Grosz, Weinstein, & Joshi, 1995), *Structure Building Framework* (e.g., Gernsbacher, 1991) and *Mental Salience Framework* (Chiarcos, 2009). For an overview of initial empirical research on visual perceptual salience see, e.g., Sridhar (1988).
2. An entity is contrastive if it is chosen from a limited set of possible entities that speaker and listener have in mind (Chafe, 1976). Contrast can co-occur with topic (i.e., contrastive topic; see Krifka (2008) for an overview of information-structural notions).
3. Bock and Warren (1985) define conceptual accessibility as the ease with which a referent`s mental representation is retrieved from memory. Prat-Sala (2000) divides conceptual accessibility into inherent accessibility (i.e., features of the referent itself like animacy) and derived accessibility (i.e., prominence features of the referent in the linguistic or non-linguistic context).

Acknowledgements

We thank Franziska Machens, Helene Killmer, Tobias Busch, and Jan Ries for assistance in material preparation and data collection.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the German Research Foundation (DFG) [grant number SFB 632] "Information structure". AS was partially funded by the German Federal Ministry of Education and Research (BMBF) [grant number 01UG1411]. The charge for publishing open access was funded by the University of Potsdam.

ORCID

Juliane Burmester http://orcid.org/0000-0001-6636-143X

References

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics*, *24*(1), 65–87.

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 29–88). Amsterdam: John Benjamins Publishing Company.

Arnold, J. E. (1998). *Reference form and discourse patterns* (*Dissertation*). Stanford University.

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*(4), 521–536.

Arnold, J. E., & Lao, S.-Y. C. (2015). Effects of psychological attention on pronoun comprehension. *Language, Cognition and Neuroscience*, *30*(7), 832–852. doi:10.1080/23273798.2015.1017511

Baayen, R. H. (2008). *Analysing linguistic data*. Cambridge: Cambridge University Press.

Bader, M., & Häussler, J. (2010). Word order in German: A corpus study. *Lingua. International Review of General Linguistics. Revue internationale De Linguistique Generale*, *120*(3), 717–762.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore, & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development*. (pp. 131–158). New York: Psychology Press.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). arXiv preprint arXiv:1506.04967.

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

Bock, J. 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). New York: Psychology Press.

Bock, K. J., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*(1), 47–67. doi:10.1016/0010-0277(85)90023-X

Bornkessel, I., & Schlesewsky, M. (2006). The role of contrast in the local licensing of scrambling in German: Evidence from online comprehension. *Journal of Germanic Linguistics*, *18*(1), 1–43.

Bosch, P., Katz, G., & Umbach, C. (2007). The Non-subject bias of German demonstrative pronouns. In M. Schwarz-Friesel, M. Consten, & M. Knees (Eds.), *Anaphors in text: Cognitive, formal and applied approaches to anaphoric reference*. (pp. 145–164). Amsterdam: John Benjamins Publishing Company.

Bosch, P., & Umbach, C. (2007). Reference determination for demonstrative pronouns. *ZAS Papers in Linguistics*, *48*, 39–51.

Bouma, G., & Hopp, H. (2007). Coreference preferences for personal pronouns in German. *ZAS Papers in Linguistics*, *48*, 53–74.

Bower, G. H., & Morrow, D. G. (1990). Mental models in narrative comprehension. *Science*, *247*(4938), 44–48.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252. Retrieved from http://www.jstor.org/stable/2984418

Burkhardt, P., & Roehm, D. (2007). Differential effects of saliency: An event-related brain potential study. *Neuroscience Letters*, *413*(2), 115–120. doi:10.1016/j.neulet.2006.11.038

Burmester, J., Spalek, K., & Wartenburger, I. (2014). Context updating during sentence comprehension: The effect of aboutness topic. *Brain & Language*, *137*, 62–76. doi:10.1016/j.bandl.2014.08.001

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 26–55). New York: Academic Press.

Chiarcos, C. (2009). *Mental salience and grammatical form: Toward a framework for salience metrics in natural language generation* (*Dissertation*). Universität Potsdam.

Colonna, S., Schimke, S., & Hemforth, B. (2012). Information structure effects on anaphora resolution in German and French: A crosslinguistic study of pronoun resolution. *Linguistics*, *50*(5), 901–1073.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45.

Cowles, W. (2007). The influence of "aboutness" on pronominal coreference. *ZAS Papers in Linguistics*, pp. 23–38.

Cowles, H. W., & Ferreira, V. S. (2012). The influence of topic status on written and spoken sentence production. *Discourse Processes*, *49*(1), 1–28. doi:10.1080/0163853X.2011.635989

Cowles, H. W., Walenski, M., & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi*, *26*(1), 3–18. doi:10.1007/s11245-006-9004-6

Crocker, M. W., Knoeferle, P., & Mayberry, M. R. (2010). Situated sentence processing: The coordinated interplay account and a neurobehavioral model. *Brain & Language*, *112*(3), 189–201. doi:10.1016/j.bandl.2009.03.004

Dahan, D., Tanenhaus, M. K., & Salverda, A. P. (2007). How visual information influences phonetically-driven saccades to pictures: Effects of preview and position in display. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eyemovements: A window on mind and brain* (pp. 471–486). Elsevier: Oxford.

Dik, S. (1997). *The theory of functional grammar*. Berlin: Mouton De Gruyter.

Flores d"Arcais, G. B. (1975). Some perceptual determinants of sentence construction. In G. B. Flores d'Arcais (Ed.), *Studies*

in perception: Festschrift for fabio metelli (pp. 344–373). Milan: Martello-Giunti.

Frederiksen, J. R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes*, *4*(4), 323–347. doi:10.1080/01638538109544525

Frey, W. (2004). A medial topic position for German. *Linguistische Berichte*, *198*, 153–196.

Fukumura, K., van Gompel, R. P. G., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *Quarterly Journal of Experimental Psychology*, *63*(9), 1700–1715. doi:10.1080/17470210903490969

Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Hove: Psychology Press.

Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. *The Psychology of Learning and Motivation*, *27*, 217–263. doi:10.1016/S0079-7421(08)60125-5

Givón, T. (Ed.). (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins Publishing Company.

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. doi:10.1016/j.jml.2007.01.007

Goodrich Smith, W., & Hudson Kam, C. L. (2015). Children's use of gesture in ambiguous pronoun interpretation. *Journal of Child Language*, *42*(3), 591–617. doi:10.1017/S0305000915000045

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., Yves von Cramon, D., & Schlesewsky, M. (2006). Linguistic prominence and broca's area: The influence of animacy as a linearization principle. *NeuroImage*, *32*(3), 1395–1402. doi:10.1016/j.neuroimage.2006.04.213

Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, *21*(2), 203–225.

Gundel, J. K. (1985). "Shared knowledge" and topicality. *Journal of Pragmatics*, *9*, 83–107.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*(2), 274–307.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English. *Journal of Linguistics*, *3*(1), 177–274.

Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, *57*(4), 596–615. doi:10.1016/j.jml.2007.01.008

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, *62*(1), 10–20. doi:10.1026/0033-3042/a000029

Hemforth, B. (1993). Kognitives parsing: Repräsentation und Verarbeitung sprachlichen Wissens (*Dissertation*). Universität Bochum.

Holle, H., Obermeier, C., Schmidt-Kassow, M., Friederici, A. D., Ward, J., & Gunter, T. C. (2012). Gesture facilitates the syntactic analysis of speech. *Frontiers in Psychology*, *3*, 74. doi:10.3389/fpsyg.2012.00074

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Järvikivi, J., van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: Contrasting the first-mention and subject-preference accounts. *Psychological Science*, *16*(4), 260–264.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*, 71–115. doi:10.1207/s15516709cog0401_4

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354.

Kaiser, E. (2006). Effects of topic and focus on salience. In C. Ebert, & C. Endriss (Eds.), *ZAS working papers in linguistics. Proceedings of Sinn und Bedeutung 10* (pp. 139–154). Berlin: ZAS.

Kaiser, E. (2011). Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus. *Language and Cognitive Processes*, *26*(10), 1625–1666. doi:10.1080/01690965.2010.523082

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, *94*(2), 113–147. doi:10.1016/j.cognition.2004.01.002

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156. doi:10.1016/S0749-596X(03)00023-8

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95–127. doi:10.1016/j.cognition.2004.03.002

Knoeferle, P., Habets, B., Crocker, M. W., & Münte, T. F. (2007). Visual scenes trigger immediate syntactic reanalysis: Evidence from ERPs during situated spoken comprehension. *Cerebral Cortex*, *18*(4), 789–795. doi:10.1093/cercor/bhm121

Knoeferle, P., & Kreysa, H. (2012). Can speaker gaze modulate syntactic structuring and thematic role assignment during spoken sentence comprehension? *Frontiers in Psychology*, *3*, 538. doi:10.3389/fpsyg.2012.00538

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

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

Lenth, R. V. (2016). Least-Squares means: The R package Lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. doi:10.18637/jss.v069.i01

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

MacWhinney, B. (1977). Starting points. *Language*, *53*, 152–168. doi:10.1145/249094.249103

Matzke, M., Mai, H., Nager, W., Rüsseler, J., & Münte, T. F. (2002). The costs of freedom: An ERP – study of non-canonical sentences. *Clinical Neurophysiology*, *113*, 844–852.

Mitchell, D. C. (1984). An evaluation of subject-paced reading tasks and other methods of investigating immediate processes in reading. In D. E. Kieras, & M. A. Just (Eds.), *New*

methods in reading comprehension research (pp. 69–89). Hillsdale, NJ: Lawrence Erlbaum Associates.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.

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*, 396. doi:10.3389/fpsyg.2011.00396

Nappa, R., & Arnold, J. E. (2014). The road to understanding is paved with the speaker's intentions: Cues to the speaker's attention and intentions affect pronoun comprehension. *Cognitive Psychology*, *70*, 58–81. doi:10.1016/j.cogpsych.2013.12.003

Nappa, R., Wessel, A., McEldoon, K. L., Gleitman, L. R., & Trueswell, J. C. (2009). Use of speaker's gaze and syntax in verb learning. *Language Learning and Development*, *5*(4), 203–234. doi:10.1080/15475440903167528

Osgood, C. E., & Bock, J. K. (1977). Salience and sentencing. Some production principles. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory* (pp. 89–140). Hillsdale, NJ: John Wiley & Sons.

Portner, P. (2007). Instructions for interpretation as separate performatives. In K. Schwabe, & S. Winkler (Eds.), *On information structure, meaning and form: Generalizations across languages* (pp. 407–425). Amsterdam: John Benjamins Publishing Company.

Prat-Sala, M. (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. doi:10.1006/jmla.1999.2668

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *1*, 53–94.

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*(8), 912–927. doi:10.1080/01690965.2013.854918

Rosengren, I. (1993). Wahlfreiheit mit Konsequenzen -Scrambling, Topikalisierung und FHG im Dienste der Informationsstrukturierung. In M. Reis (Ed.), *Wortstellung und informationsstruktur* (pp. 251–312). Tübingen: Niemeyer.

Schumacher, P. B. (2014). Content and context in incremental processing: "the ham sandwich" revisited. *Philosophical Studies*, *168*(1), 151–165. doi:10.1007/s11098-013-0179-6

Sgall, P., Hajičová, E., & Panevová, J. (1986). *The meaning of the sentence in its semantic and pragmatic aspects*. Dordrecht: Reidel Publishing, Academia.

Speyer, A. (2008). German vorfeld-filling as constraint interaction. In A. Benz, & P. Kühnlein (Eds.), *Constraints in discourse* (pp. 267–290). Amsterdam: John Benjamins Publishing Co.

Sridhar, S. N. (1988). *Cognition and sentence production: A cross-linguistic study*. New York: Springer-Verlag.

Staudte, M., & Crocker, M. W. (2011). Investigating joint attention mechanisms through spoken human–robot interaction. *Cognition*, *120*(2), 268–291. doi:10.1016/j.cognition.2011.05.005

Staudte, M., Crocker, M. W., Heloir, A., & Kipp, M. (2014). The influence of speaker gaze on listener comprehension: Contrasting visual versus intentional accounts. *Cognition*, *133*(1), 317–328. doi:10.1016/j.cognition.2014.06.003

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*(4), 519–548. doi:10.1080/01690969408402130

Stevenson, R., Nelson, A., & Stenning, K. (1995). The role of parallelism in strategies of pronoun comprehension. *Language and Speech*, *38*(4), 393–418. doi:10.1177/002383099503800404

Tannenbaum, P. H., & Williams, F. (1968). Generation of active and passive sentences as a function of subject and object focus. *Journal of Verbal Learning and Verbal Behavior*, *7*, 246–250. doi:10.1016/S0022-5371(68)80197-5

Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In J. Nuyts, & E. Pederson (Eds.), *Language, culture & cognition: Vol. 1. Language and conceptualisation* (pp. 162–189). Cambridge: Cambridge University Press.

van den Noort, M., Bosch, P., Haverkort, M., & Hugdahl, K. (2008). A standard computerized version of the reading span test in different languages. *European Journal of Psychological Assessment*, *24*(1), 35–42. doi:10.1027/1015-5759.24.1.35

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S–plus*. New York: Springer.

Vogels, J., Krahmer, E., & Maes, A. (2013). Who is where referred to how, and why? The influence of visual saliency on referent accessibility in spoken language production. *Language and Cognitive Processes*, *28*(9), 1323–1349. doi:10.1080/01690965.2012.682072

Weskott, T., Hoernig, R., Fanselow, G., & Kliegl, R. (2011). Contextual licensing of marked OVS word order in German. *Linguistische Berichte*, *225*(16), 3–18.

Zhang, L., & Knoeferle, P. (2012). Visual context effects on thematic role assignment in children versus adults: Evidence from Eye tracking in German. In N. Miyake (Ed.), *Proceedings of the 34th annual conference of the cognitive science society* (pp. 2593–2598). Japan: Cognitive Science Society.