LINGUISTIC AND VISUAL SALIENCE IN SENTENCE COMPREHENSION

EVIDENCE FROM BEHAVIOURAL AND ELECTROPHYSIOLOGICAL STUDIES

by

JULIANE BURMESTER

Doctoral dissertation
submitted to the Faculty of Human Sciences
at the University of Potsdam
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Cognitive Science

Supervisiors: Prof. Dr. Isabell Wartenburger, Prof. Dr. Katharina Spalek

Reviewers: Prof. Dr. Isabell Wartenburger, Prof. Dr. Pia Knoeferle

Submitted on 1st July 2019

Defended on 18th December 2019

Published online at the
Institutional Repository of the University of Potsdam:
https://doi.org/10.25932/publishup-44315
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-443155

Meiner Oma Liselotte

Abstract

Interlocutors typically link their utterances to the discourse environment and enrich communication by linguistic (e.g., information packaging) and extra-linguistic (e.g., eye gaze, gestures) means to optimize information transfer. Psycholinguistic studies underline that –for meaning computation– listeners profit from linguistic and visual cues that draw their focus of attention to salient information (e.g., Bornkessel, Schlesewsky, & Friederici, 2003; Knoeferle & Kreysa, 2012; Kristensen, Engberg-Pedersen, Højlund Nielsen, & Wallentin, 2013; Staudte, Crocker, Heloir, & Kipp, 2014; Weskott, Hoernig, Fanselow, & Kliegl, 2011). This dissertation is the first work that examines how *linguistic* compared to *visual salience* cues influence sentence comprehension using the very same experimental paradigms and materials, that is, German subject-before-object (SO) and object-before-subject (OS) sentences, across the two cue modalities. Linguistic salience was induced by indicating a referent as the *aboutness topic*. Visual salience was induced by *implicit* (i.e., unconscious) or *explicit* (i.e., shared*)* manipulations of listeners' attention to a depicted referent. In Study 1, a selective, facilitative impact of linguistic salience on the context-sensitive OS word order was found using offline comprehensibility judgments. More precisely, during online sentence processing, this impact was characterised by a reduced sentence-initial *Late Positivity* which reflects reduced processing costs for updating the current mental representation of discourse (e.g., Schumacher, 2014). This facilitative impact of linguistic salience was not replicated by means of an implicit visual cue (Study 2) shown to modulate word order preferences during sentence production (Gleitman, January, Nappa, & Trueswell, 2007). However, a gaze shift to a depicted referent as an indicator of shared attention eased sentence-initial processing similar to linguistic salience as revealed by reduced reading times (Study 3). Yet, this cue did *not* modulate the strong subject-antecedent preference during later pronoun resolution like linguistic salience. Taken together, these findings suggest a significant impact of linguistic *and* visual salience cues on sentence comprehension, which substantiates that *both* the information delivered via language and via the visual environment is integrated into the mental representation of the discourse; but, the way how visual salience is induced is crucial to its impact.

Acknowledgements

Meinen tiefst empfundenen Dank spreche ich für die Betreuerin meiner Doktorarbeit Prof. Isabell Wartenburger aus. DANKE, für die wertvollen wissenschaftlichen Impulse und stets zuverlässige, konstruktive, kontinuierliche, ermutigende, und aufrichtige Unterstützung! Durch sie bin ich während des Studiums mit den neurokognitiven Methoden zur Erforschung der Sprachverarbeitung zusammengewachsen und entwickelte daraus meinen eigenen Forschergeist. Meine gemeinsame Forschungszeit mit Maria Etzien bei Prof. Lyndsey Nickels, Dr. Lisi Beyersman und Dr. Britta Biedermann am Macquarie Centre for Cognitve Science in Sydney hat dann den letzten nötigen Spirit zum Schreiben meiner Doktorarbeit entfacht.

Ganz besonders danke ich auch Prof. Katharina Spalek für ihre konstruktive Kritik und ihr Feedback zu meinen Manuskripten, die sie damit stets lohnenswert bereichert hat. Ich danke beiden, Isabell und Katharina dafür, dass die Studien dieser Dissertation im Rahmen des Projektes C7 „The effect of linguistic and non-linguistic contextual information on the use of information-structural alternatives in language production and language processing“ im SFB 632 „Informationsstruktur“ entstehen konnten. Ebenso danke ich der Deutschen Forschungsgesellschaft (DFG) als finanzieller Mittelgeber meiner fruchtbaren Forschungsumgebung im Rahmen des SFBs am Department Linguistik der Universität Potsdam. Ich danke den Studentischen MitarbeiterInnnen Franziska Machens, Vivien Radke, und Tobias Busch für ihre zuverlässige Unterstützung. Ich bin meinen folgenden ProjektleiterInnen Prof. Julia Festman und Prof. Jan Lonneman sehr dankbar für das Vertrauen und die Freiheit meine Zeit für projekt- und dissertationsrelevante Aufgaben selbstbestimmt einzuteilen.

Mein bewundernswerter Dank geht an Jan Ries, der mit seinen effizienten und zuverlässigen Lösungen für komplexe Aufgaben der Visualisierung und Programmierung und die kritischen und smarten Anregungen am Rande stets gewinnbringend zu meiner Arbeit beigetragen hat. Grandios! Danke.

Ich bedanke mich bei allen TeilnehmerInnen an den durchgeführten Studien.

Von Herzen danke ich meinen zeitweisen und langjährigen BegleiterInnen in und ferner Golm mit denen ich außerordentliche gegenseitige inhaltliche und moralische Unterstützung erleben durfte, aus denen Freundschaften erwachsenen sind; insbesondere Romy Räling, Antje Sauermann, Silke Telkemeyer, Maja Stegenwallner-Schütz, Karsten Manske, Tina Marusch, Julia Holzgrefe-Lang, Franziska Machens, Helene Killmer, Laura Bos, Cornelia Heyde, McLoddy Kadyamusuma. Ihr stets offenes Ohr und Verständnis wertschätze ich sehr. Ein explizites Dankeschön geht für ihre „letzten“ wichtigen Anmerkungen an Maja Stegenwallner-Schütz, Antje Sauermann, Romy Räling und Tina Marusch.

Meinen innigsten Dank spreche ich meinen Kindern und Axel aus, für die Liebe und Energie die ich jeden Tag mit ihnen leben darf, ihre herzliche Unterstützung und ihre unendliche Nachsicht, wenn ich nochmal „ganz kurz“ was am Computer schreiben musste. Herzlichst danke ich auch meinen Eltern, meiner Schwester, meinen Omas und Opa, die mir ein reichliches Maß an Neugierde, Fleiß und Durchhaltevermögen mit auf den Weg gegeben haben. Ich danke meiner Mama und meinem Papa, dass sie mich zum Studieren mit meinem Fahrrad im kleinen Golf in Potsdam abgesetzt habt, auch wenn es ihnen schwer viel. Und ich bin in liebevoller Dankbarkeit für Oma Tine und Opa Dietmar, die mir wertvolle Zeit für das Schreiben der Arbeit geschenkt haben, durch ihre Fürsorge für unsere Kinder. Ich schätze mich glücklich, in einer so liebevollen Umgebung zu wachsen und Menschen zu wissen, deren Unterstützung ich mir sicher sein darf. Ich wertschätze mich selbst dafür, immer wieder Ideen, Motivation und Zeit gefunden zu haben, diese Arbeit in dieser Form einzureichen.

Contents

List of figures

List of tables

List of abbreviations

In addition to the standard abbreviations for Latin words, units of measurement, and those approved for the use in APA publications (American Psychological Association, 2010), the following abbreviations are used in this dissertation:

ACC	accusative case
DP	determiner phrase
DP1	first determiner phrase
DP2	second determiner phrase
eADM	extended Augmented Dependency Model
EEG	electroencephalography
ERP	event-related potential
Exp.	Experiment
ISI	inter-stimulus interval
I	Items
ISPH	Information Structure Processing Hypothesis
LAN	Left anterior negativity
Log	logarithmic
LPC	Late positive component
NOM	nominative case
OS	object-before-subject
P	Participants
PP	prepositional phrase
ROI	region of interest
SDM	Syntax-Discourse Model
SO	subject-before-object
TW	time window
V	verb

List of original journal articles

The present dissertation is based on the following peer-reviewed publications and manuscripts:

Study 1:

Juliane Burmester, Katharina Spalek, and Isabell Wartenburger (2014). Context updating during sentence comprehension: The effect of aboutness topic. *Brain and Language, 137*: 62–76. DOI: 10.1016/j.bandl.2014.08.001

Study 2:

Juliane Burmester, Katharina Spalek, and Isabell Wartenburger (2019). Visual attention-capture cue in depicted scenes fails to modulate online sentence processing. *Dialogue and Discourse, 10*(2): 79-104. DOI: 10.5087/dad. 2019.204.

Study 3:

Juliane Burmester, Antje Sauermann, Katharina Spalek, and Isabell Wartenburger (2018). Sensitivity to salience. *Language, Cognition and Neuroscience, 33*(6): 784–801. DOI: 10.1080/23273798.2018.1428758

I. SYNOPSIS

Chapter 1.

General introduction

Everyday communication is usually a fast flow of information between the interlocutors and the process of mutual understanding happens almost as fast. However, sentence comprehension does not only comprise the understanding of linguistic elements within a sentence (e.g., single words), but is a complex interplay of information processing beyond the word and sentence-level and beyond the processing of purely linguistic input. Instead, several cognitive operations need to work in concert: selecting the relevant parts of information delivered through different input modalities, linking this information with personal knowledge and shared beliefs, as well as processing several pieces of information at once (e.g., co-speech gestures), among the many other processes all associated with the individual cognitive capacities of the listener[1]. To optimise information transfer, listeners profit from information that is *salient* in discourse. In this respect, information structure deals with information packaging preferences at the prosodic, lexical, or syntactic level, adapted to the temporary communicative needs of the interlocutors (Chafe, 1976). For instance, speakers typically place that part of information about which they intend to expand the listeners' knowledge (i.e., the *topic*) in sentence-initial position (e.g., Gundel, 1985). In addition, speakers often use reduced lexical forms such as pronouns to refer to linguistically highly salient information (e.g., the topic and/or the preceding grammatical subject)

[1] In the present dissertation, the term *listener* comprises listeners of auditory linguistic stimuli *and* readers of written linguistic stimuli.

(e.g., Arnold, 1998). Such information packaging mechanisms have been explained with reference to discourse models (e.g., Garnham, 2001). These discourse models are mental meaning representations into which speakers and listeners integrate all relevant upcoming linguistic *and* perceptual input (e.g., Bower & Morrow, 1990; Gernsbacher, 1991; Grosz & Sidner, 1986; Johnson-Laird, 1980; Van Dijk & Kintsch, 1983; Zwaan, 2004). In these models, the information status of discourse referents has been proposed to be represented in terms of degrees of *accessibility* or activation states, whereas *topic* is attributed to a very high accessibility degree (e.g., *Accessibility Theory* by Ariel, 2001).

The impact of diverse purely linguistic cues on sentence comprehension has extensively been proven by studies using self-paced reading, acceptability judgements, electrophysiological, and neuroimaging methods (Bornkessel et al., 2003, Bornkessel & Schlesewsky, 2006a, Cowles & Ferreira, 2012; Cowles, Kluender, Kutas, & Polinsky, 2007; Kristensen et al., 2013; Kristensen, Engberg-Pedersen, & Wallentin, 2014; Schumacher & Hung, 2012; Weskott et al., 2011). However, during face-to-face communication, speakers often refer to the relevant part of information in the visual environment by looking at it or by using co-referential gestures to draw listeners' attention. Importantly, a growing branch of psycholinguistic research highlights the role of non-linguistic, visual salience cues on sentence and pronoun comprehension, and production, respectively (e.g., Gleitman et al., 2007; Holle et al., 2012; Knoeferle & Kreysa, 2012; Nappa & Arnold, 2014; Vogels, Krahmer, & Maes, 2013).

The present dissertation aims to shed light on the parallels of linguistic vs. visual salience-induced effects on a referent's accessibility in the mental representation of discourse, which should be reflected in the impact of these cues on the comprehension process. For this purpose, the linguistic concept of topic proved to be a fruitful starting point for comparing its impact on sentence comprehension with visual salience effects. Hence, linguistic salience is induced by means of topic-hood of a referent. Visual salience is induced by means of implicit or explicit manipulation of the listener's attention to a depicted referent. Taken together, the three empirical studies of this dissertation are the first that examine how linguistic compared to visual salience cues

influence sentence comprehension using the very same experimental paradigms and materials, that is, German subject-before-object (SO) and object-before-subject (OS) sentences, across the two cue modalities. This allows me to tackle the underlying theoretical research question, if the information status of topic-hood is grounded in the linguistic domain or if it is linked to underlying cognitive mechanisms such as selective attention. I used electrophysiological and/or different behavioural methods to test fine-grained dissociable effects of linguistic salience cues (Study 1 and 3) and two types of visual salience cues (implicit cue: Study 2; explicit cue: Study 3) on sentence comprehension.

In the following sections of Chapter 1, I will review the underlying theoretical and empirical background as well as methodological considerations for answering the current research question as outlined in Chapter 2. In Chapter 3, I briefly summarise the three empirical studies and their major results, before I combine the findings of each study into a joint discussion and conclusion. The published, original, peer-reviewed articles of Study 1, 2, and 3 are presented in Part II, following this Synopsis.

1.1 Models of sentence and discourse comprehension

Researchers in the field of linguistics and cognitive sciences have proposed theoretical and neurocognitive models in an attempt to account for the complex processes of meaning computation during sentence and discourse comprehension. Some accounts of sentence comprehension explicitly outline how the sentence itself is processed with respect to its linguistic units, that is, the analysis of prosodic, lexical, syntactic, and/or thematic information, whereas their predictions about serial, modular, or interactive processing mechanisms for meaning computation are based on behavioural (e.g., Frazier & Fodor, 1978; Marslen-Wilson & Tyler, 1980) or neurocognitive evidence (e.g., Bornkessel & Schlesewsky, 2006b; Friederici, 2002).

The research question of this dissertation encompasses a broader perspective beyond the sentence-level. Discourse-level comprehension involves multiple sources of information, ranging from interlocutors' personal knowledge, their cognitive capacities, their shared knowledge, and their visual-spatial surrounding. In this respect,

frameworks of *mental models* (also called situational or discourse models) propose that speakers and listeners (each) build a non-linguistic meaning representation in which all relevant information, provided not only by the linguistic input but by the whole situation, is held active (e.g., Bower & Morrow, 1990; Gernsbacher, 1991; Grosz & Sidner, 1986; Johnson-Laird, 1980; MacWhinney, 1977; Van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Further incoming input is incrementally integrated into these meaning representations. This leads to constant updating of the hitherto built representation. In linguistic research, these mental representations are assumed to vary in their degree of accessibility (or activation state) and hence serve as the basis for explaining information structural phenomena such as the use and interpretation of (reduced) referential expressions such as pronouns (see Garnham (2001) and Ariel (2001) as well as Section 1.3.2). Importantly, these models of sentence and discourse processing are inspired by theories assuming that cognition is grounded in multimodal perception and action, hence the comprehension process is not uni-modal and restricted to linguistic input, but is influenced by perceptual and/or action representations (e.g., *Immersed Experiencer Framework* by Zwaan, 2004). Accordingly, the *Coordinated Interplay Account* (Crocker, Knoeferle, & Mayberrry, 2010) highlights the role of visual attention during sentence comprehension by assuming closely temporally interconnected processing mechanisms for the integration of visual and linguistic information in listeners' mental representations: Following this account, sentence comprehension comprises three processing stages: 1) *incremental sentence interpretation*, 2) *utterance-mediated visual attention*, and 3) *visual scene integration,* which are based on a growing body of eye-tracking and event-related potential (ERP) research. The first stage resembles aspects of traditional sentence processing accounts assuming incremental, predictive, and integrative processing mechanisms that are responsible for representing sentence fragments in meaningful representations (e.g., Friederici, 2002; Marslen-Wilson & Tyler, 1980). However, the second and third processing stage account for listeners' rapid utterance-mediated attention shifts to predictable or anticipated referents in visual scenes and the high relevance of visual scene information for meaning computation (e.g., Knoeferle, Crocker, Scheepers, & Pickering, 2005; Knoeferle, Habets, Crocker, & Münte, 2007; Knoeferle & Kreysa, 2012).

Hence, this account provides a profound theoretical explanation for the close coupling of visual attentional and linguistic processing mechanisms during meaning computation. However, none of the existing models specifies the role of linguistic compared to visual information with respect to their strength of impact on listener's mental representations, and hence, on sentence and discourse comprehension.

The *Syntax-Discourse Model* (Schumacher & Hung, 2012) as a neurocognitive account of discourse processing assumes that meaning computation highly depends on the surrounding context, which includes sentential context, situational context, world knowledge, and common ground amongst other sources of information (Schumacher, 2014). The model characterises two temporally distinct processing mechanisms assumed to compute meaning during sentence comprehension. The first mechanism of meaning computation (called *discourse linking*) is attributed to the generation of expectations (or predictions) for upcoming words (Schumacher, 2014). The second mechanism (called *discourse updating*) is attributed to processes of updating the mental model by means of inferential reasoning. According to this model, these underlying mechanisms of meaning computation are reflected in the neural correlate of the N400 evoked around 400 ms after stimulus onset sensitive to processes of discourse linking, whereas the P600 or Late Positivity evoked around 600 ms after stimulus onset is sensitive to processes of discourse updating. In support of the Syntax-Discourse Model previous research also interpreted the Late Positivity in terms of reflecting integration and updating processes of mental representations (e.g., Delogu, Drenhaus, & Crocker, 2018; Kaan, Dallas, & Barkley, 2007; or within the neuro-computational model of language comprehension by Brouwer, Crocker, Venhuizen, & Hoeks, 2017). Hence, these ERP components characterised within the Syntax-Discourse Model allow identifying the underlying discourse-level mechanisms of sentence processing, in particular, of sentences whose processing difficulties depend on the preceding context. In this respect, German sentences with its relative flexible word order have shown to offer a promising test case for examining the impact of a preceding discourse context on the processing of sentences with varying (especially, non-canonical) word order (e.g., Meng, Bader, & Bayer, 1999; Weskott et al., 2011).

1.2 Word order in German

German is a language with a strong subject-first-ordering preference (Hemforth, 1993), but a reordering of sentence constituents is possible without changing their grammatical function because morphological features of case marking at noun phrases allow grammatical role assignment: Nominative case (NOM) is assigned to the subject, whereas accusative case is assigned to the object. Note that for masculine but not for feminine noun phrases, subjects and objects are unambiguously differentiated via case marking. In the preferred, canonical word order, the subject precedes the object (SO sentences, see example 1.1)[2], whereas in the non-canonical word order the object precedes the subject (OS sentences, see example 1.2).

1.1 SO: Der Hamster wäscht den Käfer.
[the$_{[NOM]}$ hamster$_{[NOM]}$]$_{subject}$ [washes]$_{verb}$ [the$_{[ACC]}$ beetle$_{[ACC]}$]$_{object}$.
`The hamster washes the beetle.´

1.2 OS: Den Hamster wäscht der Käfer.
[the$_{[ACC]}$ hamster$_{[ACC]}$]$_{object}$ [washes]$_{verb}$ [the$_{[NOM]}$ beetle $_{[NOM]}$]$_{subject}$.
`The hamster, the beetle washes.´

For the purpose of the present dissertation, I primarily focus on these sentence-initial (also called *prefield*) word order variations of subject and object, although, sentence-medial word order variation (so-called *scrambling*) is also possible and follows similar contextual constraints (e.g., Bornkessel & Schlesewsky, 2006a; Weskott et al., 2011).

Word order in German is influenced by different ordering principles such as the information status of discourse referents, that is, the given-before-new-ordering or the topic-comment structure (e.g., Jacobs, 2001; Weber & Müller, 2004). Hence, the German prefield tends to be occupied by the topic, if no scene-setting or contrastive element competes for this position (e.g., Frey, 2004a; Speyer, 2008; also see Section 1.3.1 for more information about the concept of topic-hood). However, SO sentences are

[2] Note that in the examples, the nouns “Hamster” [hamster] and “Käfer” [beetle] lack overt case affixes, but the determiner is overtly case marked.

felicitous and easy to process if presented without a specific, preceding context, that is, "out of the blue" (Bader & Meng, 1999, p. 129) or as an answer to a so-called *wide focus* question (*"What happened?"* or *"What did he say?"* as in e.g., Bader & Meng, 1999; Skopeteas & Féry, 2007; Stolterfoht, Friederici, Alter, & Steube, 2007) (cf. example sentence 1.1 for a compatible answer)[3]. In contrast, this is not compatible with the less frequent OS word order (e.g., Bader & Häussler, 2010), that elicits processing difficulties if presented without a felicitous context as revealed by ERPs (e.g., Bornkessel et al., 2003; Matzke, Mai, Nager, Rüsseler, & Münte, 2002; Rösler, Pechmann, Streb, Röder, & Henninghausen, 1998; Schlesewsky, Bornkessel, & Frisch, 2003). But, multiple evidence suggests that processing costs for OS sentences *de*crease or are even eliminated if presented within a *felicitous* linguistic or visual context (e.g., Holle et al., 2012; Knoeferle et al., 2007; Knoeferle & Kreysa, 2012; Stolterfoht et al., 2007; Weskott et al., 2011, see Section 1.4 and 1.5 for a detailed overview).

1.3 Terminological issues

In the field of information structure, the terms *topic*, *accessibility* and *salience* are commonly mentioned along the row with *prominence* and *high activation* depending on the different theoretical frameworks (e.g., *Accessibility Theory* by Ariel, 2001; *Prague's school* by Sgall, Hajičová, & Panevová, 1986; *Centering Theory* by Grosz, Weinstein, & Joshi, 1995; *Structure Building Framework* by Gernsbacher, 1991, *Mental Salience Framework* by Chiarcos, 2009; see Arnold, Kaiser, Kahn, & Kim, 2013 for an overview). In the following sections, I will give an overview of different definitions in order to clearly delineate these terms from one another in the present dissertation.

1.3.1 Topic

Information structural notions such as *topic* lack a uniform definition due to different theoretical frameworks (e.g., Chafe, 1976; Frey, 2007; Jacobs, 2001; Lambrecht, 1994). According to a cognitive understanding of topic, topic activates the listener's mental representation right at the beginning of a sentence (e.g., Portner, 2007) while

[3] Focus describes the new, contrastive, or informative part of an utterance (Féry & Krifka, 2008).

performing "the anchoring role to the previous discourse or to the listener's mental world" (Vallduvi & Engdahl, 1996, p. 465). Topics are associated with a high degree of accessibility, activation state, or cognitive status within mental representations (e.g., Ariel, 1988; Givón, 1983; Gundel, 1985). According to the pragmatic understanding of topic by Reinhart (1981), a *sentence topic* or *aboutness topic* is defined as "what the sentence is about" (also by e.g., Kuno, 1972; Dik, 1978; as cited in Reinhart (1981), p. 58). This is equivalent to the understanding of topic as that part of information about which the speaker intends to increase the listener's knowledge (Gundel, 1985). Krifka (2008) also points to the communicative intention of information structural concepts such as topic, namely by adopting a communication model based on a continuous change of the common ground comprising information shared by the interlocutors (e.g., Stalnaker, 1974). For topic, Krifka (2008) proposes a file-card structure of information storage whereby topic constitutes that entity in the content of the common ground under which the following information (i.e., the *comment*) is stored. *Comment* is understood as a complementary notion of topic, describing what is said about the topic (Gundel, 1985). The example sentence in 1.4 illustrates a typical topic-comment order in German. Studies aiming at eliciting typical aboutness-topic-structures (across different languages), presented speakers with depicted visual scenes while asking the context question *"What about [the topic]?"* (Skopeteas et al., 2006, see example in 1.3).

1.3 Was ist mit dem Hamster?
"What about the hamster?"

1.4 Der Hamster wäscht den Käfer.
[the hamster]$_{\text{topic}}$ [washes the beetle]$_{\text{comment}}$.
"The hamster washes the beetle."

Hence, at the syntactic level, speakers typically mention topics in sentence-initial position (e.g., Hockett, 1958; Jacobs, 2001; Lambrecht, 1994). Because in many languages this position is occupied by the subject, topics are often subjects. Still, in languages with relative flexible word order such as German, topics can be realised as grammatical objects, namely sentence-initially in OS sentences in order to make

sentences correspond to the topic-first ordering preference (cf. example 1.5). At the lexical level, topics can be expressed by definite noun phrases (e.g., *"the hamster"* in example 1.4 or 1.5) or by anaphors such as personal pronouns (e.g., *"he"* in example 1.6) (Ariel, 1988) referring to the topic as the antecedent referent in the preceding discourse (for instance, following a context question like 1.3 or sentences like 1.4 and 1.5).

1.5 Den Hamster wäscht der Käfer.
[the hamster]$_{\text{topic}}$ [washes the beetle]$_{\text{comment}}$.
"The hamster, the beetle washes."

1.6 Er freut sich auf das Picknick.
[He]$_{\text{topic}}$ [is looking forward to the picnic]$_{\text{comment}}$.
"It is looking forward to the picnic."

In addition to these syntactic and lexical features, topics are identifiable by their discourse-level features, which, however, are not obligatory. For instance, topic typically refers to information which is *discourse-given* or *old*, but in some cases, topic can comprise *new* (i.e., first time introduced) information (e.g., Frey, 2004b; Krifka, 2008; Reinhart, 1981). In contrast to *new* information, *given* information is assumed to be known by the listener due to explicit mention or inference from the previous context (e.g., Chafe, 1976; Givón, 1983; Gundel, 1988; Schwarzschild, 1999). If topics comprise information that is contrastive and hence, indicates an alternative aboutness topic or an alternative set of referents, the notion of *contrastive topics* is used (Büring, 2011; Krifka, 2008). In German sentence production, contrastive topics are preferably placed sentence-initially, similar to the positioning of sentence or aboutness topics (Frey, 2004a; Skopeteas & Féry, 2007).

Within this dissertation, the term *topic*[4] *is understood as a pragmatic concept of aboutness* (Reinhart, 1981) *activating the listener's mental representation sentence-initially* (Portner, 2007). I broaden this understanding of topic by also including its possible cognitive origin which builds the bridge to human attention, as proposed by

[4] In this dissertation the terms *topic*, *topic-hood*, or *topic status* are used interchangeably and refer to the same concept as described here.

Tomlin (1997): Accordingly, topic reflects the conceptual representation of a referent attracting the speaker's attention at the moment of utterance formulation, which – adapted to the present work on comprehension– is *a referent attracting the listener's attention at the moment of sentence comprehension*. This understanding of topic lays the foundation for my research question, that is, to test, if topic is grounded in the linguistic domain, or if manipulations of listeners' visual attention induce a similar impact on the accessibility of referents in mental representations.

1.3.2 Accessibility

Across the diverse disciplines of cognitive sciences, *accessibility* is a broad term used to describe the level of activation of mentally represented information, which is necessary to retrieve for further processing steps and that is linked to cognitive processes such as attention and memory. In linguistic research, accessibility has been linked to the information status of discourse referents and events represented within mental models (e.g., Garnham, 2001). Therein, the degree of accessibility varies along a continuum (*Accessibility Theory* by Ariel, 2001). Relevant incoming information leads to the integration of information and the dynamic process of updating the mental model (Schumacher & Hung, 2012). For instance, due to a context like in example 1.3, the "hamster" is indicated as the topic that increases its accessibility in the mental model. If this referent is repeated or recently mentioned in sentences such as 1.5 and 1.6, its accessibility is further increased (e.g., Ariel, 1988, or see Arnold, 1998 for a detailed overview of factors influencing accessibility).

At the syntactic level, highly accessible referents such as topics are typically attributed to the sentence-initial position (see also Section 1.3.1), which listener's retrieve ahead further upcoming information (Gernsbacher, 1991; Levelt, 1989; MacWhinney, 1977). This designated position for highly accessible referents has a profound impact on mental meaning computation across different frameworks, namely as the starting point or the foundation to which subsequent information is attached (Gernsbacher & Hargreaves, 1988; Gernsbacher, 1991; MacWhinney, 1977). This assumption is grounded in findings showing that first-mentioned referents are easier to access and more often correctly remembered (e.g., Gernsbacher, 1991 and references

therein). At the lexical level, accessibility is reflected as follows: The more accessible a mental representation is, the more reduced is its referential expression. Hence, reduced lexical forms such as pronouns refer to highly accessible information (e.g., topics, as outlined in Section 1.3.1), whereas full lexical noun phrases, which are much more explicit forms, refer to less accessible, for instance, new information (Ariel, 1988; Garnham, 2001; Gundel, Hedberg, & Zacharski, 1993).

Moreover, accessibility is related to the predictability of a referent in a discourse which is determined by, for instance, the presence of other potential referents (Ariel, 1988; Givón, 1988). Essentially, empirical evidence for this is provided by studies adding visual scene information to production tasks: Pronoun use is reduced and naming latencies are increased if several referents are visually depicted, although a pronoun could have clearly identified the antecedent referent (Arnold & Griffin, 2007; Fukumura, Van Gompel, & Pickering, 2010). These results were explained by a competition of accessibility between several referents, and hence can be interpreted as evidence in favour of visual context effects on the accessibility of mentally represented discourse referents. These studies suggests that accessibility modulations are not limited to linguistic cues, but maybe tightly coupled to mechanisms of visual attention. In this respect, Arnold and Lao (2015) suggest that either, *shared attention* of interlocutors or listeners' *egocentric attention* influences the accessibility of mental representations. The shared-attention hypothesis is supported by evidence showing that speakers adapt their utterances (e.g., the type of referential expression) to the presumed focus of attention of the listener (e.g., Chafe, 1976; Gundel et al., 1993) and by the proven effect of eye gaze and pointing gestures on pronoun comprehension (e.g., Nappa & Arnold, 2014). Or, alternatively the impact of attention on the accessibility of mental representations is limited to listeners' *egocentric attention*, and even not the shared aspect with speakers, which is supported by evidence showing that listeners follow their egocentric attention shifts during visual scene perception (e.g., Arnold & Lao, 2015; Bower & Morrow, 1990; Nappa & Arnold, 2014).

In contrast to these studies that define accessibility at the discourse-level, other production studies link the term accessibility to the lemma or conceptual level. Bock

and Warren (1985) describe *conceptual accessibility*[5] as the retrieval of a referent's mental representation from memory or "representing potential referents in thought" (p. 47). As a result, highly accessible referents are earlier retrieved from memory, and hence, are assigned the most prominent syntactic role (Levelt, 1989), that is, the sentence-initial subject, as these studies are based on English as a subject-prominent language (e.g., Gleitman et al., 2007; Myachykov, Thompson, Garrod, & Scheepers, 2012). However, it seems that the understanding and effects of *mental* and *conceptual* accessibility are not so different as they share common features: A referent's accessibility degree is 1) increased by visual information the speaker is attending to (e.g., Ariel, 1988; Arnold & Lao, 2015; Gleitman et al., 2007), 2) increased by semantic and thematic referent-related features such as animacy or agent-hood (e.g., Ariel, 1988; Levelt, 1989; Prat-Sala & Branigan, 2000), and 3) reflected in the linear ordering within a sentence with highly accessible referents mentioned sentence-initially. In this dissertation, I consider the different origins of both accessibility concepts, but use them as parallel evidence. For present purposes, the term *accessibility is used to refer to the activation state of a referent's (non-linguistic) mental representation. Hence, highly accessible referents refer to more easily retrievable mental representations of this referent* (e.g., Arnold, 2010). *Moreover, mentally represented referents become highly accessible through their specific salience features in discourse, which can result, for instance, from linguistic or visual properties which draw the addressee's focus of attention to that referent* (Vogels, 2014).

1.3.3 Salience

Essentially, for the purpose of this dissertation, the term *salience refers to information (i.e., referents) currently in the focus of attention of the addressee*[6] (e.g., Arnold et al., 2013), *which hence increases the accessibility of that information in the mental representation.* This definition functions closely with Tomlin's (1997) cognitive

[5] Note that conceptual and mental accessibility refer to non-linguistic mental representations that operate independent of the word itself but refer to its information (e.g., Arnold, 2010). In contrast, *lexical accessibility* is defined as the ease with which *linguistic* representations, that is, wordforms can be retrieved from memory (Bock & Warren, 1985, p. 52).

[6] The term *addressee* is used to describe someone to whom certain, linguistic and/or visual stimuli is presented (i.e., a listener, reader, or observer).

understanding of topic as that referent currently in the focus of attention of the speaker, or listener when adapting this definition to the comprehension process. Hence, in this dissertation, the information structural concept topic is used as an experimental implementation of *linguistic salience*. *Visual salience* is experimentally implemented by *implicit* and *explicit* visual cues that draw the addressee's attention to a depicted referent. Implicit visual cues, draw the addressee's attention to a depicted referent unconsciously, that is, below the threshold of perception. Explicit visual cues draw the addressee's attention above the threshold of perception, for instance, *attentionally* by arrows or *intentionally* by eye gaze. My interpretation of salience is based on its discourse-level definition as "the sum of factors that influence the degree of accessibility of an entity in the mental model" (Burkhardt & Roehm, 2007, p. 115). Hence, I suggest that the salience of entities in discourse, be it due to linguistic (e.g., topic-hood, subject status) or visual means (e.g., implicit or explicit manipulations of listener's attention), increases their accessibility degree in listeners' mental representation (e.g., Ariel, 1988; Arnold, 2010; Arnold & Lao, 2015; Givón, 1988; Gundel et al., 1993).

Indicators of linguistic salience can be derived from the factors associated with increased degrees of accessibility as outlined in Section 1.3.2, for instance, sentence-initial mention and referring to it by pronouns. However, the graded impact of different indicators is debated. For instance, topic, contrastive topic, and focus are information statuses all explicitly associated with salient information in discourse (Ariel, 1988; Arnold, 1998; Arnold et al., 2013; Frey, 2006; Kaiser, 2006). As there is widespread consensus that pronouns refer to the most salient referent in the discourse, studies on pronoun resolution most importantly contribute to the search of salience factors. In this respect, pronouns have been found to refer to topic referents (Ariel, 1988; Bosch & Umbach, 2007; Colonna, Schimke, & Hemforth, 2014; Cowles, 2007; Gundel et al., 1993; Rohde & Kehler, 2014), but syntactic prominence in terms of the grammatical subject and first mention most strongly contributes to a referent's salience (e.g., Arnold, 1998; Frederiksen, 1981; Frey, 2007; Järvikivi, Van Gompel, Hyönä, & Bertram, 2005; Kaiser, 2006; for German: Bosch, Katz, & Umbach, 2007; Bouma & Hopp, 2007; Colonna, Schimke, & Hemforth, 2012). Moreover, animate referent's are more salient than

inanimate referents, and with respect to theta roles, agents of the action are more salient than patients, which again is reflected in word order and the type of referring expression used by speakers (e.g., Ariel, 1988; Grewe et al., 2006; Jackendoff, 1990; Kaiser, 2011).

In addition to these linguistic salience cues, a body of psycholinguistic research used multiple visual cues indicating depicted referents (e.g., agents and patients of a visual scene) as salient; for instance, via attentional cues such as arrows, intentional cues such as eye gaze and pointing gestures, or foregrounding of a referent. These cues –similarly to linguistic cues– modulated the comprehension and production of different word orders (Gleitman et al., 2007; Myachykov et al., 2012; Sridhar, 1988; Tomlin, 1995; Vogels et al., 2013), and preferences of pronoun interpretation and use (Arnold & Lao, 2015; Nappa & Arnold, 2014). Besides these linguistic and visual salience cues, Osgood and Bock (1977) suggest a more complex interaction of multiple salience principles, which are related to, for instance, the speaker's intentions reflected in the tendency of producing highly motivated information first. However, again this salience factor, manifests itself in the stable principle of mentioning information in the speaker's focus of attention sentence-initially. In the following two sections, I will outline the empirical evidence for linguistic and visual salience effects on sentence comprehension and production in more detail.

1.4 Empirical evidence for linguistic salience effects

Numerous psycholinguistic studies using behavioural and electrophysiological methods have contributed to our understanding of mental representations that are built during language comprehension (Hasson & Giora, 2006). In the following, I present empirical evidence concerning the impact of different linguistic salience cues, with a special focus on the effects of a discourse referent's information status on sentence comprehension and pronoun resolution in German.

1.4.1 Behavioural findings

The influence of linguistic contexts on German sentence comprehension and pronoun resolution has been studied by behavioural methods such as acceptability judgements, self-paced-reading, and visual-world eye-tracking. In Meng et al. (1999), German SO and OS sentences were presented with preceding questions either indicating a subject-focus (i.e., *"Who did the grandpa visit?"*), object-focus (i.e., *"Who visited the grandpa?"*), or establishing a neutral context (i.e., *"What did Fritz tell you?"*). Readers were asked to judge whether or not these question-answer pairs represent a felicitous pair. Judgements showed a preference for the sentence-medial SO compared to OS word order if preceded by a neutral context. By contrast, object-focus questions increased acceptability judgements for OS sentences such that OS sentences were as acceptable as SO sentences. Additionally, these findings from acceptability judgements were supported by reduced reading times of OS sentences following the object-focus question compared to the neutral context. For SO sentences, the preceding neutral or subject-focus context did not modulate the already high acceptability judgments and fast reading times, indicating a significant processing advantage of SO sentences compared to context-sensitive OS sentences. However, for the German prefield, sentence-initial word order variations showed a different preference: Topic-hood, but not focus, of the sentence-initial object led to similarly high acceptability judgements of OS and SO sentences (Weskott, Stolterfoht, Bornkessel, & Schlesewsky, 2004). Also for the German prefield, increased acceptability judgements and reduced reading times for OS sentences were reported following a contrastive (whole-part relation) compared to a neutral context (Weskott et al., 2011). For SO sentences the contrastive context did not elicit differences in the acceptability compared to the neutral context; but – compared to OS sentences– SO sentences following the contrastive context were rated as less acceptable. These findings underline the strong role of a contrastive feature between the preceding discourse context and the following sentence-initial object. Similar findings of acceptability judgements are reported for sentence-medial word order variations, but ERPs during sentence processing still revealed higher difficulties for the OS word order embedded in contrastive contexts (Bornkessel & Schlesewsky, 2006a). Instead, a corrective context did not increase acceptability judgements for OS,

but reduced local syntactic processing costs at the object-first position. These findings suggest a benefit of using both online and offline methods to capture the impact of context-induced effects on sentence-comprehension at different processing levels: local (online) processing vs. global (offline) processing (Bornkessel & Schlesewsky, 2006a).

Furthermore, behavioural evidence for linguistic salience effects on pronoun comprehension underline the strong impact of a referent's grammatical subject status in contrast to first-mention and topic status on antecedent choice preferences for German personal pronouns (Bouma & Hopp, 2007). By contrast, based on reading times for pronoun resolution and the distribution of pronoun antecedents in German corpora, Bosch and Umbach (2007) report effects of linguistic salience in form of the grammatical subject and discourse topic. Accordingly, as revealed by antecedent preferences during listening to sentences with pronouns, Colonna et al. (2012) showed that pronouns were resolved in favour of the subject by default; but despite this subject-antecedent preference, topic-hood of the object increased its accessibility as indicated by an increase of antecedent choices in favour of the object.

In sum, evidence from acceptability judgements and reading times emphasises the sensitivity of OS sentences with respect to different linguistic salience cues. In particular, results for the German prefield underline the importance of topic-hood and the contrastive feature of the sentence-initial object indicated by the preceding context (Weskott et al., 2004; Weskott et al., 2011). Together with evidence from pronoun resolution, the mentioned behavioural findings speak in favour of a strong subject bias with respect to a referent's salience, which might be modulated by the information status of discourse referents such as topic-hood.

1.4.2 Electrophysiological findings

A growing body of electrophysiological studies aimed to get a deeper insight in the underlying mechanism that the behavioural findings point to. Several studies in accordance with the Syntax-Discourse Model (Schumacher & Hung, 2012) show that the N400 and the Late Positivity are modulated by the information status of discourse referents in the preceding linguistic context (see Section 1.1 for more detailed

information about this model). For instance, discourse linking processes during processing of German sentences, were reflected in a more pronounced N400 for discourse-new compared to discourse-given referents (Burkhardt, 2006), or for shifted topics compared to continuous topics (Schumacher & Hung, 2012). Moreover, an increase of memory capacity demands for drawing inferences and the underlying process of discourse updating led to a more pronounced P600 or Late Positivity, for instance for discourse-new (compared to inferable) referents (Burkhardt, 2006; Kaan et al., 2007). Furthermore, facilitative effects of preceding linguistic contexts indicating the object as given or as a corrective focus were reflected in a reduced so-called *Scrambling negativity* evoked around 400 ms time-locked to the processing of sentence-medial OS compared to SO word orders (Bornkessel et al., 2003; Bornkessel & Schlesewsky, 2006a). Moreover, cognitive costs for integrating a focussed sentence-initial subject or object into the current discourse model were reflected in an early positivity around 300 ms, which Bornkessel et al. (2003) interpreted in terms of the P3b ERP component, that has been reported to correlate with the expectedness of presented stimuli.

In sum, these electrophysiological studies demonstrate that discourse-level effects on sentence processing manifest themselves in well-known language-related ERP components such as the P3b, N400, and P600 (or Late Positivity) (see Van Berkum, 2012 for an overview) or even eliminate online processing difficulties of the (sentence-medial) OS compared to SO word order in German. Hence, ERPs are well-suited for examining salience-induced effects on the underlying neurocognitive mechanisms of meaning computation.

1.5 Empirical evidence for visual salience effects

1.5.1 Behavioural findings

Evidence from both sentence comprehension and sentence production demonstrate a significant impact of several visual salience cues on the addressee's attention during the perception of visual scenes. Various types of visual salience cues such as a human's or robot's eye gaze or abstract arrows to visually depicted information influenced

sentence comprehension in German (e.g., Knoeferle & Kreysa, 2012; Staudte & Crocker, 2011). For instance, listeners observed a speaker's gaze shift to one of several depicted referents during the processing of SO and OS sentences (Knoeferle & Kreysa, 2012). This gaze-induced manipulation of listeners' attention led speakers to anticipate the upcoming referent, which eased sentence comprehension, shown by faster reaction times for thematic-role verifications.

During pronoun resolution, visual salience induced by speaker gaze aligned with a pointing gesture to a depicted referent modulated the strong preference of interpreting the first-mentioned referent as co-referential with a subsequent pronoun (first-mention preference), as the cued referent increased its likelihood of being chosen as the antecedent (e.g., Nappa & Arnold, 2014). Eye-gaze and gestures are strong social-communicative cues signalling shared-attention of speaker and listener. Hence, these cues can be described as *intentional* cues, as they do not just capture the addressee's attention like abstract cues in the form of arrows (henceforth: *attentional* cues). This distinction between intentional and attentional cues is relevant to explain differences in the effect of visual salience cues on pronoun comprehension, and thus perhaps also on the accessibility of discourse referents in the mental model. In fact, findings by Nappa and Arnold (2014) support this view because only intentional –but *not* attentional– cues had the power to modulate listeners' strong first-mention preference in the search for a potential antecedent during pronoun resolution. More strikingly, after participants were instructed that the attentional cue is intentionally created by the speaker, the effects were similar to the once induced by the "natural" intentional (gaze and pointing) cue. In a similar paradigm, visual attentional cues modulated listeners' trial-initial attention shifts to visually depicted referents, but these cues only marginally influenced the strong first-mention and subject preference during pronoun resolution (Arnold & Lao, 2015). However, these findings were interpreted in favour of an impact of attentional cues on the accessibility of referents in the discourse, yet linguistic cues such as first-mention or subject status show a stronger impact (Arnold & Lao, 2015). In contrast to the power of visual cues with an intentional feature reported for pronoun resolution (Nappa & Arnold, 2014), Staudte et al. (2014) showed that the qualitative benefit of both an intentional (i.e., speaker gaze) and purely

attentional (i.e., arrow) salience cue is almost equal to listeners' comprehension (as revealed by reaction times for sentence verifications); and also elicits similar attention shifts during visual scene inspection (as revealed by eye-tracking).

Essential evidence in favour of a strong impact of several attentional visual cues is provided by research of sentence production. Implicit (unconscious) or explicit salience cues to depicted referents draw speakers' attention in such a way that the cued referent was earlier accessible than other (not cued) referents, which led speakers to mention the cued referent sentence-initially, so that the speaker's word order preferences were modulated in subsequent picture descriptions (Gleitman et al., 2007; Myachykov et al., 2012; Tomlin, 1997; Vogels et al., 2013).

Besides these effects of visual salience cues, psycholinguistic studies, in which listeners were presented with visual scenes but without additionally manipulating their attention, report effects of visual scene integration during sentence processing: For instance, Sedivy et al. (1999) showed that the accessibility degree of activated discourse entities is modulated by visually depicted contrastive features of the same entity (a yellow vs. pink comb) as indicated by shorter eye movement latencies compared to depicted entities without this contrast. Moreover, depicted events of the sentential referents eased comprehension and disambiguation of German SO and OS sentences (Knoeferle et al., 2005). In summary, a large number of visual world eye-tracking studies convincingly point to the influence of visual scenes and salience cues on sentence processing and pronoun resolution.

1.5.2 Electrophysiological findings

Effects of visual scene information –not to mention effects of visual salience cues– on sentence comprehension have been rarely investigated by measuring ERPs. However, the following two studies point to the effects of visual scene and salience cues on sentence comprehension that are similar to those ERP modulations elicited by linguistic contexts (see Section 1.4.2): A concurrently presented event visually depicting *who* is performing the action with *whom* (without any visual *salience* cues) triggered immediate (verb-mediated) syntactic reanalysis as reflected in the P600 during the

processing of locally structurally ambiguous German OS compared to SO sentences. Interestingly, the P600 showed a similar scalp topography and latency independent of the modality of the disambiguating cue, that is, *visually* depicted events or *linguistic* cues (i.e., case marking at the noun phrase) (Knoeferle et al., 2007). Moreover, processing difficulties for unambiguous OS compared to SO sentences were levelled off by the presence of these visual scenes as reflected in similar ERPs for both word orders.

Analogously to the behavioural evidence in favour of the crucial role of the intentional component of visual salience cues (Nappa & Arnold, 2014), Holle et al. (2012) showed that an intentional, speech-aligned (beat) gesture (i.e., a short hand movement) emphasising the subject of the sentences facilitated comprehension of ambiguous SO and OS sentences in German. Listeners immediately integrated the gesture-induced emphasis into processes of syntactic structure assignment reflected by reduced processing costs in terms of the P600 for the disambiguation towards the OS word order. By contrast, an explicit purely visual attentional cue (i.e., a moving point following the same trajectory as the beat gesture) did *not* ease comprehension of the same sentential structures, which hence, underlines the importance of the intentional relevance of visual salience cues.

In sum, based on these electrophysiological findings it can be suggested that visual scene and intentional salience cues are directly integrated into mechanisms of sentence comprehension, and moreover, that similar mechanisms might mediate the comprehension of both visual and linguistic contexts (Cohn, Paczynski, Jackendoff, Holcomb, & Kuperberg, 2012).

1.6 Methodological background

Across the studies of this dissertation, I used different psycholinguistic methods to test the effects of linguistic and visual salience cues on sentence comprehension with respect to readers' behavioural responses and electrophysiological processing. Behavioural methods measure participants' response after processing the critical stimulus, that is, after word, phrase, sentence, or discourse processing. Offline behavioural methods such as acceptability judgements and antecedent choice tasks

provide information about participants' conscious decision that might involve meta-linguistic processes (Weskott et al., 2011). Instead, self-paced reading is a behavioural method that allows online (i.e., real-time) investigations of sentence processing as reading times capture the time (in milliseconds) that participants need to read a word or phrase. Nevertheless, online methods such as electroencephalography (EEG) allow investigating the comprehension process "as it unfolds" (e.g., Van Berkum, 2004, p. 232), and thus are useful for gaining an insight into more unconscious and automatised mental processing stages (Hasson & Giora, 2006).

In the three empirical studies of this dissertation, I combine both offline and online methods for the following two reasons: On the one hand, I aim to capture fine-grained context-induced differences by manipulations of linguistic and implicit visual salience cues on subsequent sentence processing by using EEG (Study 1: Experiment 2; Study 2). On the other hand, it is necessary to develop reliable interpretations of results acquired by online methods. To this end, it is beneficial to collect participants' conscious decisions either within the same experimental paradigm (Study 3: self-paced reading and antecedent choices) or by comparing the effects of the identical stimulus material on ERPs and behavioural judgements (Study 1, Experiment 1 vs. 2). All of these methods, applied or adapted in this dissertation, have been used to examine linguistic or visual context effects on the comprehension of German SO and OS sentences (see Section 1.4 and 1.5).

1.6.1 Methodological considerations for the behavioural methods

In the following I will present a brief methodological description of acceptability judgements, self-paced reading, and the antecedent choice task before motivating the use of these methods in the studies of the present dissertation.

Acceptability judgements are a behavioural measure used to enquire the acceptability, grammaticality or suitability of sentences in a given linguistic context. For instance, in studies comparing word order preferences of German sentences with respect to different linguistic contexts, participants were asked to judge the acceptability of sentences in terms of a binary decision (e.g., felicitous vs. infelicitous)

per button press within a time window of 2 seconds (called *speeded* judgements in Bader & Meng, 1999; Meng et al., 1999). In Study 1 (see Chapter 5.2, Experiment 1) of this dissertation, I used judgements as a first step to capture readers' comprehensibility of stories containing either a topic or neutral linguistic context followed by an SO or OS sentence, respectively.

Self-paced reading is a common psycholinguistic method for investigating online (i.e., real-time) sentence processing in adults (Jegerski, 2014), which has been used in numerous studies interested in effects of preceding discourse-level information (e.g., Brown, Savova, & Gibson, 2012; Kristensen, Engberg-Pedersen, & Poulsen, 2014; Weskott et al., 2011). During self-paced reading participants read word- or phrase-wise presented sentences at their own pace by consecutive button presses. Researchers take reading times as an indicator of processing difficulty during the comprehension of written stimuli that have been shown to be particularly useful for examining differential effects of preceding contexts on sentence processing (Hasson & Giora, 2006). For instance, Van Berkum, Brown, Zwitserlood, Kooijman, and Hagoort (2005) show that readers make use of information from the wider discourse in order to anticipate specific upcoming words: reading times were reduced for discourse-predictable nouns, which was also supported by a reduced N400 effect in the ERPs. Note that reading times for a particular word are sometimes confounded by spillover effects reflecting a delayed effect on processing which originated by prior stimuli (e.g., Mitchell, 2004).

Pronoun reading times have been used to empirically measure the degree of referent accessibility (Hasson & Giora, 2006). For instance, shorter reading times for pronouns are interpreted as indicating an ease of reference to a discourse referent in the mental model (e.g., McKoon, Ward, & Ratcliff, 1993). During pronoun resolution, listeners identify the antecedent referent by matching the accessibility, as well as the semantic and grammatical features of all potential discourse referents against the semantic and grammatical features of the pronoun (McKoon et al., 1993 and references therein). But note that reading times for pronouns leave a caveat as they reflect the time for referential processing, but without guaranteeing to measure successful pronoun resolution –and not a *failure* in pronoun resolution– which might become true in cases

of more than one potential referent. Hence, using exclusively reading times to measure the degree of a referent's accessibility does not constitute a reliable method for identifying processing times of successful pronoun resolution as pronouns might be left unresolved (McKoon et al., 1993).

Instead, an additional *antecedent choice* or recognition task (e.g., Gernsbacher & Hargreaves, 1988), in which readers show how the pronoun was finally resolved, is needed to complement results of pronoun reading times. For instance, in a recognition task participants had to decide for or against a presented antecedent referent (i.e., target vs. distractor) (McKoon et al., 1993). In an alternative task design, that is usually conducted combined with visual world eye-tracking studies, participants are simultaneously presented with a visual scene depicting at least two potential referents, while listening to a pronoun which is embedded in sentences not favouring one of them semantically. Hence, this pronoun can be called an *ambiguous* pronoun as it could potentially refer to multiple referents in the depicted context. However, due to its linguistic salience (i.e., elicited by topic vs. focus, Colonna et al., 2014) or visual salience (i.e., elicited by different visual cue types, Nappa & Arnold, 2014) the pronoun might preferably be resolved with respect to one of these referents (the antecedent). In order to measure modulations of referent accessibility by the preceding linguistic or visual salience cue in Study 3 of the present dissertation we measured self-paced reading times during reading the ambiguous pronoun (e.g., *"He"* in *"He is already looking forward to the picnic.")* and additionally adapted an *antecedent choice task*: In this antecedent choice task we presented participants with a follow-up comprehension question about the depicted story (e.g., *"Who is already looking forward to the picnic?"*) whereupon participants had to choose one out of three visually depicted referents as the antecedent of the pronoun via mouse-click at the respective referent.

1.6.2 Event-related potentials during sentence and discourse processing

Electrophysiological methods such as electroencephalography (EEG) enable us to tackle the rapid and incremental operations of complex processes on sentence comprehension within milliseconds (e.g., Van Berkum, 2004, 2009). The EEG is a non-invasive method measuring the spontaneous electrical activity of the brain by means of

electrodes at the scalp. Event-related potentials (ERPs) index small voltage changes in response to linguistic or other sensory stimuli which evoke different patterns concerning their polarity, topography, latency, and amplitude (see e.g., Coulson, 2006 for an overview concerning the method and characteristics of different ERP components, and Van Berkum, 2012 more specifically concerning discourse processing). Within neurocognitive accounts of sentence and discourse processing (e.g., Bornkessel & Schlesewsky, 2006b; Schumacher & Hung, 2012) well-known ERP components such as the N400 and P600 have been interpreted as reflecting temporally distinct linguistic and cognitive aspects of neural information processing. While some accounts interpret these ERP components as a function of specific linguistic processing steps (e.g., N400 for semantic integration vs. P600 for syntactic integration processes as in the *auditory sentence processing model* by Friederici, 2002), other accounts favour a more general framework of the human brain's information processing across different domains (see e.g., Bornkessel-Schlesewsky & Schlesewsky, 2019 for a discussion of the N400 as reflecting predictive processing).

The N400, a negativity evoked around 400 ms after stimulus onset, is commonly attributed to semantic violations during processing of any meaningful linguistic stimuli (i.e., words or sentences) and even non-linguistic stimuli (i.e., depicted objects or visual scenes) (Cohn et al., 2012 or see Kutas & Federmeier, 2011 for a review of the N400 literature). The N400 is more pronounced for words and pictures that are less congruent, probable, or plausible in the semantic or discourse context of an unfolding sentence, and hence are harder to retrieve or integrate into the mental representation that is already activated by the preceding information (e.g., Bornkessel-Schlesewsky & Schumacher, 2016; Kutas & Hillyard, 1980; Van Berkum, 2008; Willems, Özyürek, & Hagoort, 2008).[7] As evidenced by discourse-dependent N400 modulations, the propositional meaning of perceived input can be retrieved from highly constraining linguistic discourse contexts (e.g., Otten & Van Berkum, 2007) or processing coherent sequential pictures (Cohn et al., 2012). In addition to the sensitivity of the N400 and the

[7] For the sake of brevity, I do not discuss if context-induced modulations of the N400 index dissociable processing costs for lexico-semantic access, preactivation (e.g., Lau, Phillips, & Poeppel, 2008), semantic integration, or expectancy of the critical word into the sentence context.

P600 to discourse-level processes, these ERP components reflect dissociable syntactic processing demands for the integration or revision of sentences presented in isolation, that is, without preceding context information (e.g., Friederici, 2002; Kaan, Harris, Gibson, & Holocomb, 2000; Osterhout & Holcomb, 1992; for German SO and OS sentences: Bornkessel, McElree, Schlesewsky, & Friederici, 2004; Frisch, Schlesewsky, Saddy, & Alpermann, 2002).

In sum, robust evidence supports the view of the N400 and P600 as discourse-sensitive neural correlates that are sensitive to information from different modalities such as linguistic and visually depicted information. Hence, the use of scalp EEG as a sensitive method to investigate differences in the nature and time course of sentence processing within the wider discourse, that is, in the presence of linguistic and visual context information.

Chapter 2.

Aims and research questions

The aim of this dissertation is to address the question of how linguistic compared to visual salience cues influence sentence comprehension. To answer this research question, I assembled theory-based, behavioural, and electrophysiological evidence with an emphasis on information structure together with psycholinguistic evidence with an emphasis on effects of visual scene information on sentence comprehension and production (as outlined in Chapter 1). Essentially, the evidence points to the information structural concept of topic, which is associated with a referent in the addressee's focus of attention, as is hence salient in discourse. In the linguistic domain, topic referents are attributed to a high accessibility degree in the mental representation of discourse, and are hence, typically mentioned sentence-initially and constitute the preferred antecedent for a subsequent pronoun (e.g., Ariel, 1988; Arnold, 1998; Jacobs, 2001). In the visual domain, visually salient referents are also attributed to an earlier or increased accessibility degree, which led to sentence-initial mention of this referent or to an increased likelihood to choose these referents as antecedents of a pronoun (e.g., Gleitman et al., 2007; Nappa & Arnold, 2014). Hence, linguistic salience by means of a referent's topic-hood and visual salience by means of visual cues to a depicted referent modulate two parallel information packaging preferences, that is, 1) salient referents occupy the sentence-initial position, and 2) salient referents are preferred antecedents for pronouns. This dissertation aims to provide empirical evidence for the underlying theoretical research question, whether and how discourse processing, in particular

processing of a discourse referent's information status (i.e., topic), is grounded in the linguistic domain or if non-linguistic (i.e., visual) cues could affect sentence processing and pronoun resolution similarly.

As a theoretical basis for this research question, I follow the framework of mental models in which the information status of discourse referents is represented in terms of graded accessibility (Garnham, 2001; McKoon et al., 1993) and into which linguistic as well as perceptual input from the visual-spatial surrounding is integrated (e.g., Johnson-Laird, 1980). Hence, mental models pose the link between the two modalities of salience, and thus, serve as the theoretical basis for deriving my hypothesis: Effects of linguistic or visual salience cues on the accessibility degree of mentally represented discourse referents should be reflected in the process of meaning computation during sentence and pronoun comprehension. In particular, I hypothesise that an *increase* of accessibility by salience cues should be reflected in processing ease during sentence-*initial* processing of the salient referent as this is the designated sentence position for highly accessible information. Additionally, as pronouns preferentially refer to the most salient referent in discourse, I hypothesise that the effects of salience on the accessibility degree of mentally represented referents should be visible in modulations of antecedent choice preferences during pronoun resolution. German sentences with their relative flexible word order offer a promising test case for examining the impact of preceding linguistic and visual cues on sentences with varying word order, especially because of the proven context-sensitivity of German non-canonical (OS) sentences (see Section 1.4). In the three studies outlined below, I use behavioural and electrophysiological methods to 1) test if and how linguistic salience (topic-hood) of a referent influences sentence processing (Study 1 and 3) and pronoun resolution (Study 3), and 2) compare this to the influence of two different types of visual salience cues (implicit, explicit) on sentence processing (Study 2 and 3) and pronoun resolution (Study 3).

More specifically, I address the following aims and research questions within each study:

1) **Study 1** (Chapter 5) examines how linguistic salience by means of a purely linguistic cue indicating the aboutness topic influences sentence comprehension in German. For this, I use a behavioural judgement task in order to detect if readers' overall comprehensibility of stories containing SO or OS sentences was modulated by the linguistic salience vs. neutral cue (Experiment 1). In addition, I aim to capture the nature and time course of salience-induced differences on subsequent sentence-initial processing of German SO and OS sentences by recording ERPs (Experiment 2). I hypothesise that effects of linguistic salience modulate underlying processes of meaning computation as reflected in discourse-level ERP components proposed by the Syntax-Discourse Model (e.g., Schumacher & Hung, 2012).

2) **Study 2** (Chapter 6) examines how visual salience by means of an implicit, purely attention-capturing, visual cue presented below the threshold of perception influences online sentence-initial processing. With this study, I aim to conceptually replicate Study 1 (linguistic salience) and to compare the influence of linguistic and visual salience on ERPs during sentence processing across the two modalities. I hypothesise that, if implicit visual cues influence the process of meaning computation, this effect should be reflected in similar underlying discourse-level processes as shown for linguistic salience.

3) **Study 3** (Chapter 7) examines how salience of a depicted referent due to a linguistic (i.e., topic) or an explicit, intentional visual cue (i.e., a virtual person's gaze shift) modulates sentence-initial processing and pronoun resolution by means of self-paced reading and of an antecedent choice task, respectively. I hypothesise that, if the process of meaning computation is influenced by linguistic and/or explicit, intentional visual salience cues, this should be reflected in latency modulations of sentence-initial reading times of the salient referent as well in modulations of antecedent choice preferences during pronoun resolution.

A short summary of the studies with its major results is presented and discussed in the following chapter.

Chapter 3.

Summary of studies and main findings

In the three studies of this dissertation, adult participants were presented with short stories about two or three animals that were going to perform a fictitious simple transitive action (e.g., fish painting each other) while effects of salience on sentence comprehension were examined either by recording ERPs (Study 1 and 2) and/or by behavioural measures (Study 1 and 3). The presented stories always started with a lead-in sentence due to which the referents were discourse-given (see Figure 3.1 for an example). The subsequent cue, either increased the salience of *one* referent (*topic/salience cue*) or indicated a wide focus on the scene (*neutral cue*) in order to prevent participants from primarily attending to (one of) the referents. Effects of these cues were investigated with respect to subsequent processing of German sentences with varying word order that described *who* was doing *what* to *whom* (i.e., the thematic role relations). Accordingly, in the target sentence either the subject was first-mentioned (i.e., in SO sentences) or the object was first-mentioned (i.e., OS sentences) (see Figure 3.1 for a simplified overview of the experimental design of Study 1 and 2). In all three studies, these two cue conditions (salience vs. neutral) and the word order of the target sentence (SO vs. OS) served as stable parameters, whereas the way the linguistic and visual cues were implemented, differed between the studies.

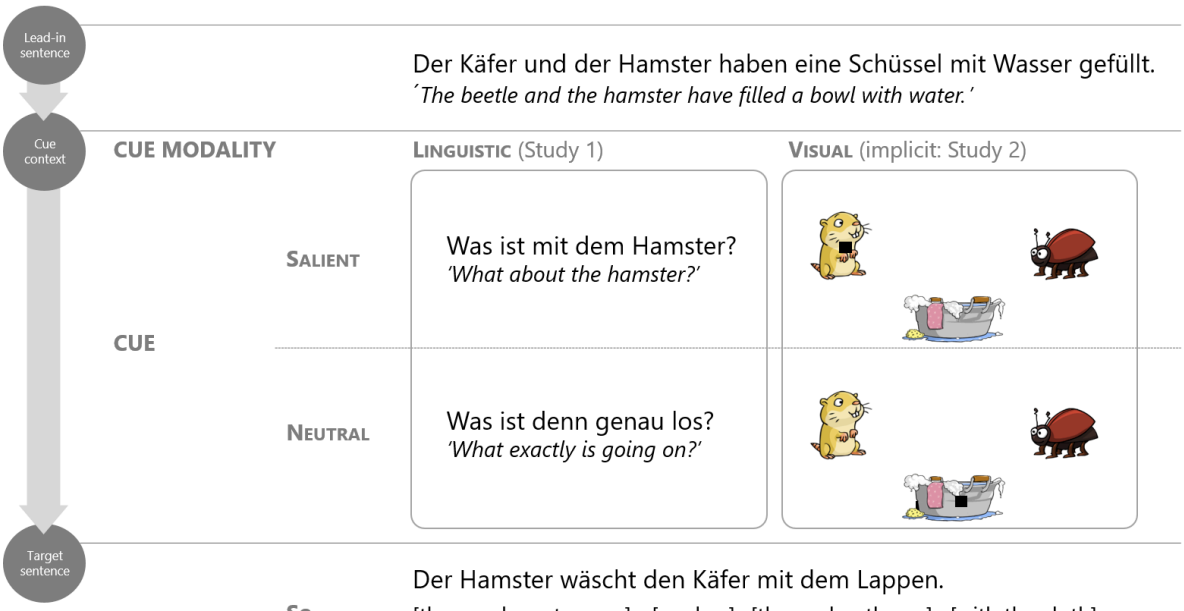

Figure 3.1: Simplified overview of experimental design of Study 1 and 2
Participants read the lead-in sentence, followed by the linguistic or visual cue context, and the target sentence presented either in the SO or OS WORD ORDER. Conditions are capitalised. Approximate English translation is written in italics. Note that in Study 2, the implicit visual cue (i.e., the black square located at the hamster (salient cue) or washing bowl (neutral cue)) was presented *before* the pictures were presented, hence, both directly followed each other.

Abbreviations: SO = subject-before-object, OS = object-before-subject, S = subject, V = verb, O = object, NOM = nominative case, ACC = accusative case, PP = prepositional phrase.

In **Study I** (Chapter 5) linguistic salience was induced by presenting readers with an aboutness topic context that explicitly mentioned the topic (e.g., *"What about the hamster?")*. The results of study 1 show that linguistic salience facilitated the comprehension of topic-first non-canonical OS sentences, but not of canonical topic-first SO sentences in comparison to a neutral cue (i.e., *"What exactly is going on?"*). The impact of linguistic salience was shown in Experiment 1 by offline comprehensibility judgments of the stories and supported by the ERP results of Experiment 2, in which I investigated the nature and time course of linguistic salience on sentence processing. The selective, facilitative effect of linguistic salience was reflected in a reduced Late Positivity around 500-700 ms at the sentence-initial topic position in OS sentences. In

line with the assumptions of the Syntax-Discourse Model (e.g., Schumacher & Hung, 2012) and previous research reporting a Late Positivity or P600 in response to discourse-level processing (e.g., Kaan et al., 2007), I attributed the Late Positivity to reduced processing costs for updating the current discourse model when the salient referent –the topic– is indicated before its processing in sentence-initial position. These findings show that the applied study design was well suited to examine effects of *linguistic* salience by means of a topic cue. Therefore, study 1 laid the foundation for the derivation of experimental designs for testing possible effects of visual salience in the two subsequent studies. Besides the selective, facilitative effect of the topic cue on OS sentence processing, the topic cue elicited an early, sentence-initial positivity peaking around 200 ms compared to the neutral cue, which was independent of word order. I interpret this early effect similarly to studies that associate early ERP modulations with basic information encoding processes (e.g., Dunn, Dunn, Languis, & Andrews, 1998). Accordingly, this early positivity follows as a result of processing repeated words in the topic, but not in the neutral condition.

In **Study 2** (chapter 6), visual salience was induced by replacing the linguistic topic cue of Study 1 by an implicit visual attentional cue presented below the threshold of perception in order to direct participants' attention to a depicted referent. The cue was a small black square presented subconsciously for a duration of 66 ms at the location of the subsequently visually depicted referent. In a previous production study by Gleitman et al. (2007) this type of cue has been shown to direct participants' attention to the cued referent, and hence, increased the cued referent's likelihood of being mentioned sentence-initially. I used ERPs analogously to Study 1 (Experiment 2) in order to be able to test fine-grained dissociable effects of the implicit visual cue on subsequent sentence-initial processing of the cued referent and to compare these results to the ones with the linguistic salience cue. However, in contrast to the facilitative impact of the linguistic salience cue on sentence-initial processing, the implicit visual cue did *not* reduce processing costs of OS sentences, neither reflected in terms of Late Positivity modulations nor of other discourse-related ERP components. Besides the absence of a cue-induced effect on sentence processing we replicated a word order (or canonicity) effect found in Study 1 and in previous studies (e.g., Matzke

et al., 2002; Rösler et al., 1998), which I interpret as a "sanity check" of the applied experimental paradigm. Moreover, although participants did not notice the presence of the cue, modulations of early sensory-evoked potentials (in terms of the N1 and P2) in response to the implicit visual cue and visual scene suggest that they indeed processed the cue; but finally, without a dissociable effect on sentence-initial processing.

In **Study 3** (Chapter 7), visual salience was induced by an explicit intentional cue, that is, a gaze-shift of a virtual person to a depicted referent, while linguistic salience was induced by a context question that indicated the topic status of a depicted referent (e.g., *"What about the left one?"*). Self-paced reading was used to measure effects of these salience cues on reading times during sentence processing as an indicator of processing effort, while pronoun reading times and an antecedent choice task was used to measure if the salience cues modulate antecedent choice preferences during later pronoun resolution. The results showed that linguistic *similar to* visual salience cues immediately led to faster sentence-initial reading times of sentences mentioning the salient referent first. Accordingly, these results indicate that the two salience cues reduced sentence-initial processing effort of both SO and OS sentences independent of modality. For later pronoun resolution, our results show the well-known subject-antecedent preference evidenced in previous studies (e.g., Bouma & Hopp, 2007), but importantly, linguistic and visual salience cues modulated this preference *differently*. Whereas visual salience did *not* modulate the subject-antecedent preference, linguistic salience cues combined with salient-first OS sentences significantly reduced the subject-antecedent preference for a subsequent pronoun in favour of choosing the salient and sentence-initially mentioned object as the antecedent. To summarise the main findings of Study 2, both linguistic and visual cues show an immediate impact on sentence-initial processing, whereas later pronoun resolution was primarily driven by linguistic cues.

In the following chapter, I will briefly interpret the results of all three studies with respect to the underlying research question of the present dissertation, namely, if and how a referent's accessibility in the mental model is modulated by inducing a referent as salient via linguistic or visual cues.

Chapter 4.

Discussion and concluding remarks

In the preceding chapter, I presented the main findings of three empirical studies that I will discuss with respect to the influence of linguistic vs. visual salience cues on sentence comprehension. Linguistic salience was induced by topic-hood (Study 1 and 3). Visual salience was induced by either implicitly or explicitly directing participants' attention to the salient referent of a story (Study 2: implicit, purely *attentional* cue, Study 3: explicit, *intentional* cue). I hypothesised that effects of linguistic and visual salience cues on the accessibility degree of a mentally represented discourse referent should be reflected in an ease of processing the cued referent in sentence-initial position as this is the designated sentence position for highly accessible referents. In addition, possible modulations of accessibility due to linguistic or visual salience cues should be visible in reduced subject-antecedent preferences during pronoun resolution as the cued referent might compete for accessibility with the salient subject referent.

The findings of Study 1 complement previous corpus-based, behavioural and electrophysiological evidence that underline the sensitivity of the German OS word order to the preceding discourse, and thus point to the significant role of the object's information status to motivate word order variation (e.g., Bader & Häussler, 2010; Bornkessel et al., 2003; Meng et al., 1999; Weskott et al., 2004; Weskott et al., 2011). In contrast, SO sentences are high frequent and easy to process even without a felicitous context (e.g., Gorrell, 2000; Weber & Müller, 2004). While the judgements revealed an overall advantage for the comprehension of SO compared to OS sentences, linguistic

salience induced by topic-hood of the upcoming object increased the likelihood of OS sentences being judged as easily comprehensible. The findings of the ERPs show that the selective, facilitative effect of the topic cue in OS sentences originates very early during processing, as reflected in the reduced Late Positivity at sentence-initial position. These ERP findings support the discourse-level interpretation of the Late Positivity as a correlate of processing costs for updating a mental model as assumed by the Syntax-Discourse Model (Schumacher & Hung, 2012). Accordingly, the indication of the topic by the preceding linguistic context reduced processing costs for updating the current discourse model because the topic with its high accessibility degree was already integrated into the hitherto built mental model when it came to its sentence-initial processing. In contrast, following the neutral cue, the discourse model had to be updated at the moment the reader encountered the sentence-initial object referent reflected in higher updating costs for the less accessible object in the OS word order.

However, in contrast to linguistic salience, the implicit, purely *attentional*, visual cue which participants could perceive only unconsciously, did *not* influence subsequent online sentence-initial processing in terms of underlying mechanisms at the discourse-level that could have been captured using ERPs (Study 2). The results of the implicit cue were interpreted in favour of a true null effect as ERPs during sentence processing were modulated in response to the varying SO and OS word order, and hence, replicated a word order effect which was also found in Study 1 and indexed by various ERP modulations in previous studies (e.g., Matzke et al., 2002; Rösler et al., 1998; Schlesewsky et al., 2003). However, a similar type of cue used by Gleitman et al. (2007) significantly increased the preference to mention the cued referent sentence-initially; but this cue did not influence the comprehension process. One reason for the null effect of this type of cue on sentence processing but the significant impact on ordering preferences in the production study by Gleitman et al. (2007) could originate from the different study designs. Such implicit type of cue might be too subliminal, and hence, its effect on accessibility too weak or short-lived to sustain and influence subsequent, hence time-delayed, sentence-initial processing in the comprehension paradigm. By contrast, in the production task by Gleitman and colleagues, speakers initiated their utterances *directly* at the moment they apprehended the visual scene. However, the

results of Study 3 show that an explicit, *intentional* visual cue by means of a gaze shift to a depicted referent significantly eased subsequent sentence-initial processing of both SO and OS sentences *similarly* to linguistic salience. Hence, these linguistic and explicit visual salience cues increased the accessibility of the sentence-initial referent such that any sentence-initial processing difficulties due to the different word order were levelled off. But, during later pronoun resolution, linguistic salience in terms of topic-hood, but *not* explicit visual salience, modulated the strong subject-antecedent preference reported by previous studies (e.g., Bouma & Hopp, 2007; Colonna et al., 2014). However, regardless of the different impact of the linguistic vs. visual salience cue on pronoun resolution, but the similar facilitative impact on sentence-initial processing, both cues modulated the comprehension process induced by visual attention shifts to a depicted referent. Therefore, the findings of Study 3 support the assumptions of the Coordinated Interplay Account (cf. processing stages of so-called *utterance-mediated attention* and *scene integration*, Crocker et al., 2010).

Taking the results of all three studies into account, I advocate the following view of salience-induced effects on the accessibility of mentally represented discourse referents: Linguistic salience is the *primary* factor for increasing a referent's mental accessibility as reflected in an ease of processing topic referents in sentence-initial position of OS sentences (Study 1) and even of canonical SO sentences (Study 3); as well as reflected in modulations of the subject-antecedent preference during later pronoun resolution (Study 3). The findings for pronoun resolution confirm previous research of dynamically interacting linguistic salience markers (topic-hood and first-mention, grammatical subject and object status) on referent accessibility (e.g., Bosch & Umbach, 2007; Bosch et al., 2007; Bouma & Hopp, 2007; Colonna et al., 2014; Kaiser, 2011). Still, although linguistic salience by means of topic-hood and first-mention in the non-canonical word order increased a referent's accessibility, linguistic salience accompanied by subject status maintains the strongest predictor for a referent's high accessibility.

Building on the findings of this dissertation and previous research comparing effects of linguistic vs. visual salience (Fukumura et al., 2010; Vogels et al., 2013) and

effects of visual salience regarding attentional vs. intentional cues (e.g., Arnold & Lao, 2015; Nappa & Arnold, 2014; Holle et al., 2012) on sentence and pronoun comprehension and production, visual salience seems to be of *subordinate* relevance for meaning computation. Moreover, compared to the power of linguistic salience, the impact of visual salience cues on meaning computation depends on multiple factors. One factor is the nature of the visual cue that is, its *attentional*, but even more importantly, *intentional* feature that manipulates listeners' attention. Given the *absent* impact of the attentional cue, but the *facilitative* impact of the *intentional* cue on sentence-initial processing (Study 2 vs. 3), our findings support existing behavioural and electrophysiological evidence in favour of the importance of the intentional feature of visual cues on listeners' comprehension of sentences and pronouns (Arnold & Lao, 2015; Nappa & Arnold, 2014; Holle et al., 2012). Hence, multiple evidence indicates that accessibility is more strongly affected by the power of social-communicative and intentional cues such as speakers' gazes and gestures that draw listeners' attention to information in the visual environment (Arnold & Lao, 2015; Nappa & Arnold, 2014) or that emphasise a specific part of information during speaking (e.g., the grammatical subject of a sentence, Holle et al., 2012). But, the findings by Staudte et al. (2014) indicate an almost equal benefit of intentional (gaze) and explicit attentional (i.e., arrow) cues to listeners' comprehension, and hence speak in favour of the feature of attention rather than that of intentionality. So maybe the explicitness of the visual cue and its relevance in the specific comprehension process is a crucial factor for its impact on the accessibility in listeners' mental representation.

However, *implicit* attentional cues lead to a less pronounced impact on accessibility during mental meaning computation, because their impact has shown to interact with various factors; that is, listener-individual factors such as listeners' egocentric preferences during perception of the visual environment (Arnold & Lao, 2015) or listeners' cognitive (i.e., attention and memory) load (see Arnold and Griffin (2007) and Arnold (2010) plus references therein, and Vogels, Krahmer, and Maes (2014) for effects on cognitive load on the production of referring expressions). And moreover, the impact of visual cues might depend on linguistic factors, that is, when linguistic information is ambiguous or less informative and the visual input is beneficial

to add information to an incomplete mental representation (Knoeferle et al., 2007). Hence processing of the implicit cue in Study 2 might be confounded by parallel processing capacities necessary for processing the immediately following depicted scene. As attention is a limited capacity (e.g., Bradley, 2007; Cowan et al., 2005), I suspect that parallel information processing mechanisms elicited by rapid, consecutively perceived input that differs with respect to its threshold of perception (implicit cue vs. depicted scene), and with respect to its modality (visual scene vs. sentence processing) pose a major challenge to listeners' cognitive capacities (see e.g., Penney (1989) for representing auditorily vs. visually presented linguistic stimuli in the verbal short-term memory). Following this line of thought, it could be speculated that listeners more strongly benefit from salience cues, if these cues are of the same modality, and congruent with each other (e.g., Jefferies, Lambon Ralph, & Baddeley, 2004). This assumption is supported by the findings of Study 3 showing an effect of linguistic salience and first-mention on antecedent choices, but *not* of visual salience and first-mention. Or explaining this difference more precisely, referent accessibility was increased by linguistic salience in the form of topic-hood, and subsequently supplemented by the further linguistic salience marker of first-mention. Hence, subject-antecedent choices were reduced as the salient and sentence-initial object competed for accessibility with the subject. By contrast, maybe the visually salient referent lost the competition of accessibility with the subject during unfolding sentence processing, and hence did not reduce subject-antecedent choices. Furthermore, it might require a higher cognitive effort to synchronise linguistic and visual information during processing (e.g., Cowan et al., 2005).

To conclude, I suggest a dynamic and a competitive feature of salience influencing the accessibility of mental representations during the ongoing multi-modal input that listeners perceive while the strength of impact of salience cues correlates with their degree of perception and with shared-attentional (intentional) aspects of speaker and listener. Hence, I suppose that implicit cues might need to reach a certain threshold of perception in order to increase a referent's accessibility in the mental model and in order to maintain accessibility to be still easier accessible during subsequent processing of upcoming input. Moreover, this view is in accordance with

the impact of *explicit* attentional cues that influenced sentence comprehension similar to intentional cues (Staudte et al., 2014). Furthermore, intentional cues such as a gaze cues or gestures worked as strong salience cues measurable in terms of an immediate facilitative effect during sentence processing (similarly to linguistic cues in Study 3 of this dissertation or as shown in the studies by Holle et al., 2012 and Knoeferle & Kreysa, 2012). But, gaze-induced increases of accessibility do not sustain at the high degree when linguistic salience cues such as subject-hood compete for accessibility; hence no effect of gaze was measurable during later pronoun resolution in Study 3.

As pointed out within a dynamic and embodied approach to cognitive processing by Bradley (2007), human perception is guided by their motivational state that directs attention to relevant objects in the world, whereas emotionally relevant stimuli are processed differently from unemotional ones. Drawing the link to the raised underlying research question of the present work, that is, if the concept of topic is cognitively grounded or specific to the linguistic domain, I draw the parallels from the findings of the studies of this dissertation to the information structural definition of topic within a communication model that is based on a continuous change of the common ground, that is, information shared by the interlocutors (e.g., Stalnaker, 1974). Hence, the findings concerning the similar impact of topic and the gaze cue on sentence processing (Study 3) are in support of a cognitively grounded view of topic. This view of topic is closely linked to dynamic changes in the high accessibility degree of topic referents within the mental representation of discourse due to linguistic *and* visual salience cues that draw attention, *especially shared attention*. This conclusion is also compatible with the approach of a successful dialogue for which interlocutors need to develop aligned mental representations (Pickering & Garrod, 2004), into which especially *shared*-attentional cues such as linguistic salience and visual intentional cues might offer a window. Finally, the findings of this dissertation support the view that accessibility modulations of mentally represented discourse referents are primarily a function of the linguistic input, but that also *non-linguistic, visual perceptual* input is integrated into these mental representations, amongst the many other sources of information (e.g., Ariel, 1988; Bower & Morrow, 1990; Gernsbacher, 1991; Grosz & Sidner, 1986; Johnson-Laird, 1980; Van Dijk & Kintsch, 1983; Zwaan, 2004).

II. ORIGINAL JOURNAL ARTICLES

Chapter 5.

Study 1:

Context updating during sentence comprehension: The effect of aboutness topic[8]

Brain and Language, 137, (2014), 62–76.

JULIANE BURMESTER[1], KATHARINA SPALEK[2], AND ISABELL WARTENBURGER[1]

[1] Cognitive Sciences, Department Linguistics, University of Potsdam, Potsdam, Germany
[2] Department of German Language and Linguistics, Humboldt-Universität zu Berlin, Berlin, Germany

Abstract

To communicate efficiently, speakers typically link their utterances to the discourse environment and adapt their utterances to the listener`s discourse representation. Information structure describes how linguistic information is packaged within a discourse to optimize information transfer. The present study investigates the nature

[8] This chapter is adapted from: Burmester, J. Spalek, K., and Wartenburger, I. (2014). Context updating during sentence comprehension: The effect of aboutness topic. *Brain and Language, 137*: 62–76. DOI: 10.1016/j.bandl.2014.08.001. This study was published using American English.

and time course of context integration (i.e., aboutness topic vs. neutral context) on the comprehension of German declarative sentences with either subject-before-object (SO) or object-before-subject (OS) word order using offline comprehensibility judgments and online event-related potentials (ERPs). Comprehensibility judgments revealed that the topic context selectively facilitated comprehension of stories containing OS (i.e., non-canonical) sentences. In the ERPs, the topic context effect was reflected in a less pronounced Late Positivity at the sentence-initial object. In line with the Syntax-Discourse Model, we argue that these context-induced effects are attributable to reduced processing costs for updating the current discourse model. The results support recent approaches of neurocognitive models of discourse processing.

Keywords: Information structure; Discourse context; Aboutness topic; Sentence processing; Word order variation; ERP; Late Positivity, Syntax-Discourse Model

5.1 Introduction

In everyday communication, we typically link our utterances to the discourse environment of the interlocutor in order to efficiently achieve our communicative objectives. Besides other factors, the speaker considers background information and feedback of the listener. Linguistic (e.g., information structure, stress) as well as extra-linguistic features (e.g., gestures, eye-gaze) are dynamically used to clarify what the utterance is about and ultimately guide the cooperative listener to the communicative intention of the speaker. It has been proposed that the listener structurally represents all relevant aspects of information (e.g., participants, events) delivered via language and perception within a mental model in which further incoming discourse information is integrated (e.g., Cowles, 2003; Johnson-Laird, 1980).

Information structure (cf. information packaging) is concerned with how information is packaged within a discourse to optimize information transfer (Chafe, 1976). In this regard the idea of efficient communication was defined by Clark and Haviland (1977) as: “The speaker tries, to the best of his ability, to make the structure of his utterances congruent with his knowledge of the listener`s mental world” (p. 4). Ordering of information at the sentence-level is thought to be influenced by information

structural concepts, such as topic-comment, given-new, or focus-background (e.g., Büring, 2011; Halliday, 1968; Krifka, 2008; Lenerz, 1977). However, these information structural concepts lack a uniform definition and depend on the field of research and respective theoretical framework. For the purposes of our study, we use the following definitions: The TOPIC of a sentence is typically understood as the information that the speaker intends to increase the listener`s knowledge (Gundel, 1985). Hence, topic is defined as what the sentence is about; COMMENT is what is said about the topic (Gundel, 1988; Reinhart, 1981; see Subsection 5.1.4 for a more detailed definition of topic). GIVEN INFORMATION constitutes information the speaker expects to be already known by the listener (e.g., Haviland & Clark, 1974); that is, information explicitly mentioned in the previous discourse or information that can be entailed by the context (e.g., Chafe, 1976; Schwarzschild, 1999). In contrast, NEW INFORMATION describes information the speaker expects to introduce to the listener in the sense of "newly activating" it in the listener`s consciousness (Chafe, 1976). FOCUS refers to the new/informative or contrastive part of an utterance. Whereas, BACKGROUND denotes less relevant information (e.g., Vallduvi & Engdahl, 1996). Experimentally, focus is often induced as contrastive focus, where the newness of the information is emphasized by its contrast to previously focused information (e.g., Jacobs, 1988). A special type of contrastive focus is corrective focus, where an assumption is explicitly corrected. These information structural concepts are thought to be realized by distinct prosodic (i.e., accenting) and/or syntactic (e.g., sentence position) phenomena (see e.g., Chafe, 1976; Féry & Krifka, 2008; Skopeteas & Fanselow, 2010; Steedman, 2000).

In the present study, we aim to investigate how a previously presented context, in particular a context introducing all characters of a fictitious scene with emphasis on one of them as the aboutness topic, affects the comprehension of a subsequent canonical (subject-before-object) or non-canonical (object-before-subject) declarative sentence in German. Before we present the two experiments (Experiment 1: offline comprehensibility judgments, Experiment 2: Event-related potentials (ERP) during online sentence processing) we first give a brief overview of German word order, the underlying neurocognitive mechanisms of sentence and discourse processing, as well as previous findings concerning information structural concepts and sentence

processing relevant to understanding the motivation and predictions of the present study design.

5.1.1 Word order in German

Word order in German is relatively flexible. Reordering of constituents within a sentence can be used to highlight the communicatively relevant part of the utterance. German has a strong subject-first preference (e.g., Gorrell, 2000), but reordering of constituents within a sentence is possible, because syntactic roles can still be assigned correctly due to morphological case marking at the respective determiner or determiner and noun. Case marking of the subject by nominative (NOM) and object by accusative (ACC) case is ambiguous for feminine, neuter, and plural noun phrases, but unambiguous for masculine singular noun phrases. The example sentences (1a, b) illustrate case marking for masculine subjects and objects in German with the finite, transitive verb in the second sentence position. (1a) depicts a canonical declarative sentence with typical subject-before-object (SO) word order. (1b) depicts a non-canonical sentence with object-before-subject (OS) word order.

(1a) Der Uhu malt den Igel.
[the$_{[NOM]}$ owl$_{[NOM]}$]$_{subject}$ [paints]$_{verb}$ [the$_{[ACC]}$ hedgehog$_{[ACC]}$]$_{object}$.
`The owl paints the hedgehog.´

(1b) Den Igel malt der Uhu.
[the$_{[ACC]}$ hedgehog$_{[ACC]}$]$_{object}$ [paints]$_{verb}$ [the$_{[NOM]}$ owl$_{[NOM]}$]$_{subject}$.
`The hedgehog, the owl paints.´

Sentences (1a) and (1b) differ in the manner of information packaging (SO vs. OS order). However, both sentences induce the same propositional representation. In isolation, the OS order (cf. example 1b) is assumed to be harder to process compared to SO (Schlesewsky, Fanselow, Kliegl, & Krems, 2000), but interestingly, context information (e.g., a preceding sentence or question) has been found to ease the processing of OS sentences (e.g., Meng et al., 1999) (see Subsection 5.1.3 for the effect of information structure on the processing of word order variation in German).

Thus, in German main clauses, subjects as well as objects can appear in the sentence-initial position before the finite verb (so called prefield). Similarly, if the verb is not in the second but in final sentence position, either the subject or object can follow the complementizer (so called middlefield)[9] (see e.g., Pittner & Berman, 2008, for an overview of the topological classification of German sentences). As commonly assumed, the OS order is derived from the basic order of SO; but, depending on the theoretical framework, different movement operations are assumed to underlie word order variation in the German pre- and middlefield (e.g., Haider & Rosengren, 1998; Lenerz, 2000; Müller, 1999; see Diedrichsen, 2008, for an alternative, movement-independent account of the German sentence topology). Bornkessel-Schlesewsky and colleagues substantiate the distinction of word order variation in the pre- and middlefield from the neuroanatomical perspective (Bornkessel-Schlesewsky, Grewe, & Schlesewsky, 2012). Whereas numerous studies reported an increased activation for OS opposed to SO within the left inferior frontal gyrus (lIFG), aboutness-based sequencing (prefield) activated anterior subregions of the lIFG, but prominence-based sequencing (middlefield) activated superior subregions of the lIFG (for a review, see Bornkessel-Schlesewsky & Schlesewsky, 2012).

Several semantic and discourse-related factors have been proposed to affect the linear order of sentential constituents (e.g., concerning thematic role, actors should precede non-actors; for a review about incremental argument interpretation during processing of transitive sentences, see Bornkessel-Schlesewsky & Schlesewsky, 2009a). Numerous studies proposed factors that crucially affect word order in the German middlefield (e.g., Bornkessel-Schlesewsky & Schlesewsky, 2009b; Choi, 1996; Lenerz, 1977; Müller, 1999; Siewierska, 1993). For the purposes of our study, the most important are findings concerning the German prefield: As attested in written corpora, SO and OS sentences predominately occur with an accusative object (Bader & Häussler, 2010). SO sentences tend to contain active verbs, whereas OS order frequently occurs with verbs lacking an agent argument (i.e., passivized ditransitive and unaccusative

[9] The deviation of SO order in the German middlefield is termed *scrambling* (i.e., OS: …, dass den Igel der Uhu malt. (…, [that]$_{[complementizer]}$ [the$_{[ACC]}$ hedgehog$_{[ACC]}$]$_{[object]}$ [the$_{[NOM]}$ owl$_{[NOM]}$]$_{[subject]}$ [paints]$_{[verb]}$.)).

verbs). Further, OS is more frequent if the object is animate and the subject inanimate (Bader & Häussler, 2010), which fits the previously proposed animacy-based ordering preferences of sentential constituents (Tomlin, 1986). In the present study, we aimed to exclude confounding effects of the listed linearization preferences in order to examine the effect of aboutness topic in the prefield of SO and OS sentences. Thus, we held the following factors constant: case of the object (accusative), verb type (active, transitive), thematic roles of subject (agent) and object (patient) as well as their animacy status (animate). Persisting differences between OS and SO word order we further considered by focusing on comparing contextual effects within the respective word order.

5.1.2 Neurocognitive models of sentence and discourse processing

Different neurocognitive models of sentence comprehension have been formulated to better understand the nature and time course of online sentence processing (e.g., the *extended Augmented Dependency Model* (eADM) by Bornkessel & Schlesewsky, 2006b; the auditory sentence processing model by Friederici, 2002). Basically, the architecture of these models is assumed to be hierarchically organized in phases that specify the steps of incremental sentence comprehension and correspond with functionally separable networks at the brain level. These processing steps have been linked to specific language-related ERP components. After the prosodic analysis, indexed by a negativity peaking around 100 ms (N100), the model of Friederici (2002) proposes three phases: Phase 1 is an initial phrase-structure-building process of the sentential constituents. In phase 2, morphosyntactic as well as semantic information is integrated (i.e., thematic role assignment), indexed for instance by the left anterior negativity (LAN) and the negativity around 400 ms (N400). Phase 3 is characterized by reanalysis and repair mechanisms as indexed by the positivity around 600 ms (P600) (Friederici, 2002). Similarly, the eADM proposes three phases of sentence comprehension: In phase 1, the phrase-structure representation is built via template-mapping. In phase 2, the arguments are interpreted with regard to their thematic and prominence relations, indexed by the N400, LAN, the P600 and/or the scrambling negativity –an ERP component that has been engendered by violations in sequencing arguments according

to prominence based hierarchies in languages allowing word order variation (e.g., accusative object precedes subject in the German middlefield (Bornkessel & Schlesewsky, 2006a; Bornkessel, Schlesewsky, & Friederici, 2002; Bornkessel et al., 2003) or in Japanese (Wolff, Schlesewsky, Hirotani, & Bornkessel-Schlesewsky, 2008). In phase 3 (*"generalized mapping"*), information structural mechanisms induced by the discourse context, world-knowledge and/or prosody are taken into account and trigger well-formedness evaluation and repair processes, indexed by Late Positivities (that have been suggested to belong to the P300 component). Hence, in this final phase, sentences are evaluated according to their acceptability with respect to the context environment (Bornkessel & Schlesewsky, 2006b). This is the processing step in which we also expect to see effects of contextual manipulation in the present study.

A more recent model of Bornkessel-Schlesewsky and Schlesewsky (2013) –the *"New dorsal-ventral stream model of sentence comprehension"*– explicitly links the eADM to underlying brain structures. This model assumes two processing streams working in parallel: The ventral stream builds the sentence-level semantic representation by time-independent computations such as identification and unification of conceptual (actor-event) schemata. The dorsal stream combines time-dependent elements and establishes the syntactic (constituent) structure by time-dependent computations such as prosodic segmentation, combination of elements into category sequences, actor identification etc. The two streams are integrated in the frontal cortex which subserves cognitive control and allows for top-down-feedback, pragmatic interpretation, conflict resolution, and builds the interface with motor cortices. Discourse linking processes are also assumed to be supported by parietal brain regions (Bornkessel-Schlesewsky & Schlesewsky, 2013).

In the present study, hypotheses are based on the *Syntax-Discourse Model* (SDM) (first introduced for pronominal-antecedent relations by Burkhardt, 2005, and extended to general discourse processing in a multi-stream-model by Schumacher & Hung, 2012, and Wang & Schumacher, 2013). The SDM focuses on mechanisms of information packaging during online sentence comprehension. Therein, currently processed information is assumed to be directly interpreted and integrated in relation

to a previously established discourse representation which is built incrementally (see also the *Information Structure Processing Hypothesis* (ISPH), by Cowles, 2003). According to this model, the N400 response is related to expectation-based discourse linking, whereas the Late Positivity is evoked by discourse updating processes such as the adding of a new discourse referent, topic shift, inferential reasoning, enrichment, and/or the modification of the established discourse representation (see Wang & Schumacher, 2013, and Schumacher, 2014, for recent reviews).

5.1.3 The effect of information structure on sentence processing

Recent research in the field of information structure has raised the question how information packaging in terms of word order variation is affected by different types of context information (e.g., Büring, 2011; Fanselow & Lenertová, 2011). So far, studies on word order variation in German have mainly focused on SO and OS sentences in the absence of context information (e.g., Bader & Häussler, 2010; Bornkessel, Zysset, Friederici, Cramon, & Schlesewsky, 2005; Hemforth, 1993; Kempen & Harbusch, 2005; Matzke et al., 2002; Rösler et al., 1998). However, context information plays an important role in licensing non-canonical word orders, as evidenced by occurrence frequency in corpora, behavioral and ERP findings.

In written corpora, OS is very rare in German main clauses (Bader & Häussler, 2010; Weber & Müller, 2004), but the frequency of OS significantly increases for certain discourse contexts: At first sight, the linear order of subject and object in German main clauses was determined by givenness (i.e., increased frequency of OS if the object was given in a previous context but the subject was discourse-new); however, more decisive are the factors definiteness and pronominalization –both highly correlated with givenness (e.g., pronouns and definite noun phrases predominantly represent given, indefinite noun phrases new information) (Weber & Müller, 2004). As these factors were not of interest in our study we ruled out any confounding effects by using given, definite, and full noun phrases.

Based on behavioral data (i.e., acceptability rating and reading time), strong contextual licensing effects for OS in German main clauses have been found if the object

was in a contrastive whole-part relation to a contextually mentioned set (partially ordered set relation according to Prince, 1998) (Weskott et al., 2011). Besides, a context question, which revealed the object as given and the subject as focused, improved judgments and reading times of scrambled OS in German embedded clauses (Meng et al., 1999).

How context information modulates underlying mechanisms of online sentence processing has previously been investigated by ERPs. ERP components commonly used to investigate language processing at the semantic and syntactic level, such as the well-established N400 (see e.g., Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008 for a review) and P600 or Late Positivity (Frisch et al., 2002; Osterhout & Holcomb, 1992), have been found to be sensitive to discourse-level processing (e.g., Bornkessel et al., 2003; Burkhardt, 2007; Cowles, Kluender et al., 2007; Hung & Schumacher, 2012; Van Berkum, 2012; Wang & Schumacher, 2013). Previous ERP studies examining context effects during sentence processing revealed an impact of givenness and focus. For instance, an early positivity around 300 ms for discourse-new focused initial objects in scrambled OS as well as subjects in SO was interpreted in terms of reflecting processes of focus integration (e.g., Bornkessel et al., 2003). Furthermore, the scrambling negativity for OS in the German middlefield was enhanced if the object was given opposed to a discourse-new object (Bornkessel et al., 2003); although based on behavioral findings givenness of the object would be expected to license OS (Meng et al., 1999). In a related study, Bornkessel and Schlesewsky (2006a) compared OS with SO sentences. Any processing difficulties in terms of the scrambling negativity for OS compared to SO disappeared if a preceding context induced a corrective focus.

Moreover, modulations of the N400 and Late Positivity have been proposed to index discourse integration processes (cf. SDM by Schumacher & Hung, 2012 and Wang & Schumacher, 2013, see also Subsection 5.1.2). The N400 –modulated by different degrees of givenness– has been attributed to processing difficulties in linking the current referent to the previous discourse: For instance, in German main clauses, the N400 was enhanced for inferable vs. given subjects in SO as well as objects in OS (Schumacher & Hung, 2012). Similarly, Wang and Schumacher (2013) investigated the

influence of topic status on sentence processing. The authors were interested in how different types of discourse contexts (given vs. inferable topic vs. contrastive new) influence sentence processing in Japanese: New vs. given information revealed an N400, but the N400 was absent if the new information was expected, due to its sentential position and the respective context. This finding supports the assumption that the N400 indicates expectation-based discourse linking rather than an effect of information status per se. Further, a Late Positivity (around 500-700 ms) has been proposed to reflect processing costs for updating and correcting the current discourse model, which was assumed to be more demanding for (contrastive) new vs. inferable vs. given (topic) referents (e.g., Schumacher & Hung, 2012; Wang & Schumacher, 2013). Similarly, in Chinese, the Late Positivity has been found to be sensitive to position-specific processing demands evoked by different types of topic (given topic/topic shift/new topic) (Hung & Schumacher, 2012): The preference that the topic position is filled by a given topic (i.e., topic continuation) over topic shift and new topic was reflected in an enhanced Late Positivity. A biphasic N400-Late Positivity pattern with enhanced amplitudes for new opposing to given information was reported for subsequent non-topic positions. Hence, discourse linking and updating evoke a biphasic N400-Late Positivity pattern (e.g., Hung & Schumacher, 2012; Wang & Schumacher, 2013). But both components have also been found independent of each other: For instance, the N400 was modulated by different degrees of givenness in the German prefield (e.g., Schumacher & Hung, 2012), and the Late Positivity was modulated by different degrees of expectation in the German middlefield (Burkhardt, 2007). Hence, the SDM assumes two independent processing streams for discourse linking (N400) and updating (Late Positivity) (e.g., Wang & Schumacher, 2013). Taken together, the ERP studies support that the impact of discourse information on sentence processing is detectable in modulations of well-known ERP components, such as the N400 and Late Positivity. In this regard, the SDM strongly contributes to understanding discourse relevant processing demands modified by previously presented context information.

To sum up, word order in German has been found to be context-sensitive: As evidenced by high frequency in corpora, high acceptability ratings/low reading times and online processing measures, SO is felicitous even without a context; but OS is

constrained by certain licensing contexts. Offline methods such as acceptability ratings have been used to decide whether a certain context licenses sentence processing on a global level, whereas online methods such as ERPs have been used to characterize the underlying mechanisms of context effects during incremental sentence processing. The relevance of using offline as well as online methods to characterize the level at which context information interacts with word order during sentence comprehension has been underlined by previous findings. As already mentioned, behavioral findings revealed given objects in scrambled OS felicitous (Meng et al., 1999), whereas ERPs still revealed a scrambling negativity during online processing (Bornkessel et al., 2003). Similarly, contrastively focused objects in scrambled OS improved offline acceptability ratings, but online a scrambling negativity reflected processing costs compared to SO (Bornkessel & Schlesewsky, 2006a). Most of the previous online and offline studies in German characterized the influence of givenness, focus or topic (operationalized by different degrees of givenness or inferability) on the processing of word order variation; but online studies on different types of topic in other languages (e.g., Hung & Schumacher, 2012; Wang & Schumacher, 2013) offer a useful starting point for the predictions of the present study (see Subsection 5.1.5). Importantly, in the present study, topic was operationalized as aboutness topic (see Subsection 5.1.4), while givenness was held constant (all referents given).

5.1.4 The information structural notion of aboutness topic

Topic or aboutness topic is an important information structural concept relevant for linguistic communication (for a review, see Frey, 2007and Jacobs, 2001). As a pragmatic phenomenon, aboutness topic has been described as the entity the sentence is about (e.g., Reinhart, 1981). Topic has been assumed to perform "the anchoring role to the previous discourse or the hearer`s mental world" (Vallduvi & Engdahl, 1996, p. 465). This is in line with the account that topic usually refers to information that is given due to a previous context (e.g., Givón, 1983; Gundel, 1988; Skopeteas et al., 2006). Accordingly, Reinhart (1981) pointed out that the sentence topic is identifiable by both the context of the utterance and the linguistic structure. At the sentence-level, Hockett (1958) differentiates between the topic as what the speaker announces first and the

comment as what is said about the topic. The definition as well as the identification of topic via linguistic features has been controversially discussed (see e.g., Lambrecht, 1994 for a discussion on the *"topic-first principle"*). For German main clauses, topic has been argued to strongly tend to occur sentence-initially (e.g., Büring, 1999; Frey, 2004b; Jacobs, 2001; Rosengren, 1993; Vallduvi & Engdahl, 1996) if this position is not occupied by a competitor (i.e., a scene-setting or contrastive element) (Speyer, 2004; 2008). Besides, as in German the prefield can be occupied by non-topics, the middlefield has been argued to be designated for topic (e.g., Frey, 2004a).

Taking into consideration the properties of topic plus the relatively flexible word order, German offers a promising starting point to examine the impact of topic context on sentence processing, especially on OS sentences. It remains an open question if a context inducing an aboutness topic status of given referents crucially facilitates the overall comprehension of OS in the prefield; and especially if this effect is immediately reflected in the online processing of OS sentences in terms of discourse updating of the current mental model.

5.1.5 The present study

The goal of the present study was to characterize if and how a discourse context indicating the aboutness topic of the upcoming sentence eases the processing of the following canonical (i.e., SO) or non-canonical sentence (i.e., OS) in German declaratives. By using fictitious stories that introduced two relevant characters and the event of the scene (discourse-given), we compared the effect of two differential mini-discourse contexts: In one condition, a topic context indicated the aboutness topic status of one character of the scene; in the other condition, a neutral context indicated a wide scope of the scene. The context question used to establish the topic status is similar to previous studies investigating aboutness topic during online sentence comprehension in other languages. However, these studies modulated givenness (Hung & Schumacher, 2012, 2014) or animacy (Wang, Schlesewsky, Philipp, & Bornkessel-Schlesewsky, 2012) at the same time. Whereas all referents of the scene were discourse-given, we aimed to characterize the effect of these two discourse contexts (topic vs. neutral context) on unambiguously case marked German declaratives with

either SO or OS word order. Therefore, two experimental methods were used: (1) An offline comprehensibility judgment task to test if the participants` judgment of overall understanding of the stories with either SO or OS target sentences is affected by the type of the preceding discourse context (Experiment 1), and (2) ERPs to test how the preceding discourse context incrementally modulates the online processing of the SO and OS target sentences (Experiment 2). Note that we compared the context effect within each word order, meaning that in both experiments the very same target sentences were compared to circumvent confounding effects of prominence-related sequencing preferences (such as grammatical or thematic role). These two methods provide crucial information about both the nature and time course of discourse organizational processes elicited by the two context types.

In German main clauses, a contextually induced aboutness topic is expected to be placed sentence-initially (e.g., Büring, 1999), whereas the neutral context does not generate such an expectation. Due to the strong subject-first preference in German (e.g., Hemforth, 1993), context information revealing all sentential constituents as given should not play a crucial role for the processing of SO sentences. But as evidenced previously, for non-canonical word order, context information plays a licensing role (e.g., Bornkessel & Schlesewsky, 2006b; Weskott et al., 2011). Hence, for Experiment 1, we predicted that stories containing SO target sentences should be judged as easily comprehensible, independent of context type; whereas for stories containing OS target sentences, the preceding topic context was expected to improve comprehensibility judgments. Based on recent ERP studies, discourse organizational processes have arguably been reflected in modulations of ERPs around 400 and 600 ms during online sentence processing (see above). Similar to offline comprehensibility judgments, we do not expect any modulations by the preceding context type during online processing of SO sentences in Experiment 2. However, if the topic context creates a felicitous discourse environment for OS sentences as measured by offline comprehensibility judgments, we expect that in these sentences differential processing costs induced by the two discourse contexts should be visible during online processing. Therefore, due to direct contextual integration of the topic into the discourse model, processing costs for updating the current mental model should require less effort compared to the

neutral context. This might be reflected in modulations of the Late Positivity as this ERP component has been proposed to reflect processing costs for updating and correcting the current discourse model (e.g., Bornkessel & Schlesewsky, 2006b; Burkhardt, 2007; Hung & Schumacher, 2012; Schumacher & Hung, 2012; Wang & Schumacher, 2013). Note that we do not expect a modulation of the N400 due to the fact that all constituents are discourse-given, and hence, the linking of unexpected discourse referents is not required.

5.2 Experiment 1

In Experiment 1, participants were presented with short fictitious stories. We conducted an offline comprehensibility judgment task to detect if the participants` judgment concerning the overall comprehensibility of stories containing either an SO or OS target sentence was affected by the preceding discourse context, a topic vs. neutral context. The individual behavioral judgment of the comprehensibility of each story was recorded.

5.2.1 Materials and methods

5.2.1.1 Participants

Twenty-eight German native speakers (19 female, *M* age 24 years, age range 20-34 years) participated in Experiment 1. Twenty-six participants were right-handed and two ambidextrous as assessed by a German version of the Edinburgh Handedness Inventory (Oldfield, 1971). None reported any neurological disorder. All had normal or corrected-to-normal vision. Participants were reimbursed or received course credits for participation.

5.2.1.2 Material

The experiment used a 2 x 2 within-subject design with the factors CONTEXT TYPE (TOPIC vs. NEUTRAL) and WORD ORDER (SO vs. OS). In accordance with previous studies investigating the interaction of context and word order, we created short fictitious discourses by means of question-answer pairs (Bornkessel et al., 2003; Meng

et al., 1999) that led to significantly increased acceptability ratings compared to non-question contexts (Bornkessel & Schlesewsky, 2006a). A set of 160 experimental trials (40 trials per condition) was constructed. Each trial consisted of a three-sentence discourse depicting a scene of two animals performing a transitive action in which both were equally plausible to be the agent or patient of the scene. All trials followed the structure shown in Table 5.1. (1) In the first sentence (lead-in) of each trial, the current scene with both animals and the instrument of the to-be-performed action was introduced. Thus, in terms of information structure, the relevant characters were discourse-given (Prince, 1981) and the action was inferable (Prince, 1992) from the instrument mentioned. The same lead-in was used for all conditions. (2) The following wh-question (i.e., context question) differed with regard to the factor CONTEXT TYPE: The context question either induced a wide scope of the scene (NEUTRAL CONTEXT) or indicated one of the two animals as the aboutness topic (TOPIC CONTEXT). (3) The third sentence (target sentence) provided a plausible answer to the preceding context question by describing the final action event of the two animals. The target sentence varied according to the factor WORD ORDER and was thus presented in SO or OS order.

Table 5.1: Sample experimental trial for each condition
Vertical bars in target sentence indicate phrase-wise presentation, approximate English translation written in italics.

Lead-in	Context question	Target sentence	Condition
Der Uhu und der Igel haben eine Staffelei im Park aufgebaut. *`The owl and the hedgehog have set up an easel in the park.´*	Was ist denn genau los? *`What exactly is going on?´*	Der Uhu | malt | den Igel | im Park. [the$_{[NOM]}$ owl$_{[NOM]}$]$_{DP1}$ [paints]$_{V}$ [the$_{[ACC]}$ hedgehog$_{[ACC]}$]$_{DP2}$ [in the park]$_{PP}$. *`The owl paints the hedgehog in the park.´*	NEUTRAL SO
	Was ist mit dem Uhu? *`What about the owl?´*		TOPIC SO
	Was ist denn genau los? *`What exactly is going on?´*	Den Uhu | malt | der Igel | im Park. [the$_{[ACC]}$ owl$_{[ACC]}$]$_{DP1}$ [paints]$_{V}$ [the$_{[NOM]}$ hedgehog$_{[NOM]}$]$_{DP2}$ [in the park]$_{PP}$. *`The owl, the hedgehog paints in the park.´*	NEUTRAL OS
	Was ist mit dem Uhu? *`What about the owl?´*		TOPIC OS

Abbreviations: NOM = nominative case, ACC = accusative case, DP1 = first determiner phrase, V = verb, DP2 = second determiner phrase, PP = prepositional phrase, SO = subject-before-object, OS = object-before-subject.

The different scenes were created based on 40 animals (monomorphemic nouns, masculine gender, 1-syllabic (n = 18) to 2-syllabic (n = 22)) and 10 actions (monomorphemic verbs, transitive, accusative-assigning, 2-syllabic) with corresponding instruments and a scene-setting prepositional phrase (e.g., in the park). Note that both grammatical and thematic roles coincided (i.e., the grammatical subject was always the agent, the grammatical object was always the patient). The critical nouns and verbs were matched for written lemma frequency, type frequency and normalized log10 familiarity values, taken from the dlex database (Heister et al., 2011). To control for position effects, each noun occurred once in each of the four conditions at the first and second noun phrase position of the target sentence. Thus, each animal served four times as the agent and four times as the patient of the target sentence, respectively, always with a different action and co-animal. In the lead-in sentence, the first and second mention of the potential agent and patient was counterbalanced across conditions. Both animals of a scene always differed in the initial phoneme. To minimize possible effects of structural priming (Scheepers & Crocker, 2004), all trials were pseudo-randomized such that maximally two consecutive trials were of the same condition or had the same word order in the target sentence. To avoid any preferences of thematic role or topic continuity (Givón, 1983) caused by the previous trial, at least five trials separated the repetition of an animal, and at least two trials the repetition of an action. Four lists of 160 trials were created such that each list contained each item only once, and across all lists each item occurred once in each condition. Each participant was presented with one of the four lists.

5.2.1.3 Procedure

Similar to judgments on acceptability (Bornkessel & Schlesewsky, 2006a) or felicity (Meng et al., 1999) of paired question-answers, we used a speeded comprehensibility judgment task, in which participants were explicitly asked to intuitively judge the comprehensibility of stories within a 2000 ms time window.

Participants were tested individually, seated in a sound-attenuated booth 90 cm away from the computer screen with a button box (Cedrus response pad model RB-830) on their lap. Written instructions about the experimental procedure were given to

participants. Participants were asked to read each story attentively and silently and judge each story as fast as possible with regard to its comprehensibility. The trials were displayed visually in the center of the screen by means of the Presentation software (version 14.1, www.neurobs.com). Each trial began by presenting of a red asterisk for 1000 ms to indicate the beginning of a new scene. Before and after the lead-in, a blank screen was displayed for 200 ms. Lead-in and context question were presented as a whole in a self-paced reading manner with a minimum reading time of 3350 ms and 1400 ms, respectively. The participant had to press a button with the left thumb for further reading. Then the target sentence was presented phrase-wise (as indicated in Table 5.1) with 500 ms for each determiner phrase (DP) and prepositional phrase (PP) and 450 ms for the verb with an ISI of 100 ms (as used in previous studies, e.g., Bornkessel et al., 2003).

After the presentation of the target sentence, the participant had to perform a binary judgment on the comprehensibility of the whole preceding story by pressing a button. The participant either pressed the right index or middle finger on the respective "thumb-up" or "thumb-down" button: Thumb-up for stories that were easily comprehensible or thumb-down for stories that were less easy to comprehend. The assignment of the response buttons to the participants` right index and middle finger was counterbalanced across participants. Before the experiment started, finger positions on the respective buttons were checked by the experimenter. The response option was depicted for 2000 ms. Participants performed three practice trials to become familiar with the procedure. The experiment was split into four blocks of 40 experimental trials. No filler trials were presented to keep the experimentation time within acceptable limits for the participant (i.e., to preserve motivation and concentration, and to prohibit movement artifacts or alpha waves in the signal of the electroencephalography (EEG) in Experiment 2). The whole experimental session lasted approximately 40 minutes including self-adjusted pauses after each block.

5.2.1.4 Data analysis

For statistical data analysis of the comprehensibility judgment task, we computed logit mixed models for the categorical judgments (easily vs. less easily comprehensible),

following Jaeger (2008). We used the statistical software R (version 2.15.2, R Core Team, 2013) with the supplied lme4 package (Bates, Mächler, & Dai, 2009) for the mixed models analysis and the ggplot2 package (Wickham, 2009) for the display of the results. To analyze the categorical judgments using logit mixed models, CONTEXT TYPE, WORD ORDER and the interaction of both were defined as fixed effects, while participants and items were defined as random effects. Fixed effects were coded as +.5/-.5 contrasts resembling traditional ANOVA analyses. Model fitting started with the most complex model (Barr, Levy, Scheepers, & Tily, 2013); that is, with the full factorial set of random effects (random slope adjustments for all fixed effects for both participants and items). In a step-wise manner, the complex model was reduced by model comparisons via log-likelihood tests (e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008). Slope adjustments were excluded if they did not improve the explanatory power of the model in comparison to the simpler model without that slope adjustment. Logit mixed models were fitted by the Laplace approximation. Estimates (*b*), standard errors (*SE*), *z*-values and the level of significance (*p*) of the final logit mixed model are reported.

5.2.2 Results

Participants showed the following mean (*M*) proportion for stories judged as easily comprehensible per condition: NEUTRAL SO: $M = 0.93$ ($SE = 0.04$), TOPIC SO: $M = 0.92$ ($SE = 0.04$), NEUTRAL OS: $M = 0.37$ ($SE = 0.05$), TOPIC OS: $M = 0.54$ ($SE = 0.05$) (see Figure 5.1).

The statistical analysis of the participants` categorical judgments of the stories revealed significant main effects of CONTEXT TYPE and WORD ORDER, and a significant interaction of CONTEXT TYPE x WORD ORDER (see Table 5.2 for statistics of the final logit mixed models).[10] Post hoc logit mixed models to resolve the interaction within each WORD ORDER revealed a significant effect of CONTEXT TYPE for stories containing OS sentences, but not for stories containing SO sentences. Thus, stories

[10] Additional statistical analyses using a 2 x 2 ANOVA for the proportions of easily comprehensible judgments revealed the same result pattern as the logit mixed models analysis on the raw data of the categorical judgments.

containing the OS target sentence were more likely to be judged as easily comprehensible if presented together with the TOPIC CONTEXT. For stories containing the SO target sentence, the probability to be judged as easily comprehensible was equally high independent of the preceding CONTEXT TYPE and significantly higher than for stories with the OS target sentence.

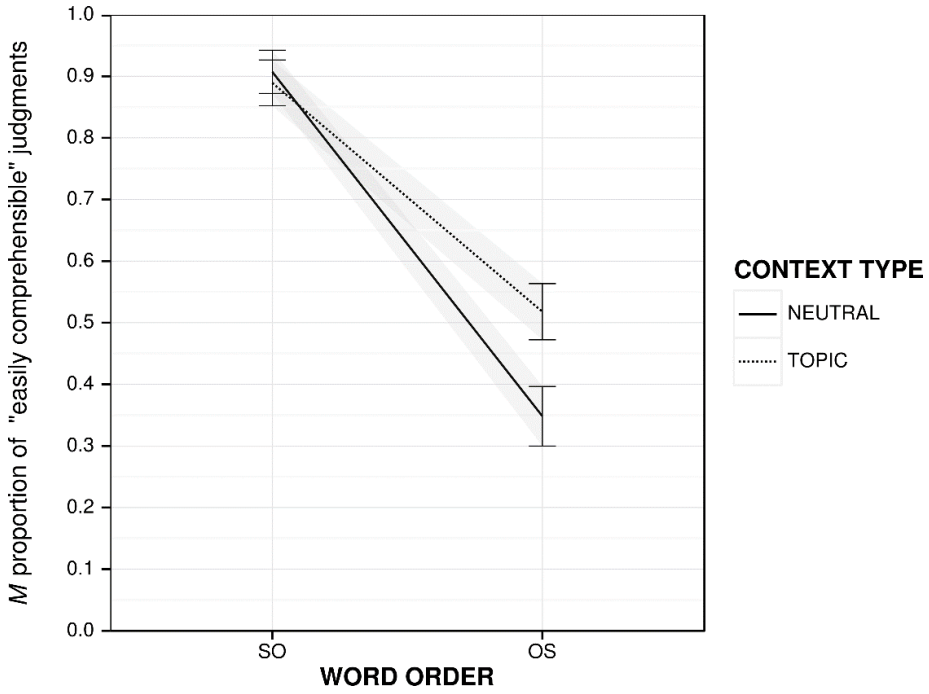

Figure 5.1: Mean *(M)* proportion of "easily comprehensible" judgements in Exp. 1
The Figure shows the *M* proportion of stories judged as "easily comprehensible" showing the effect of CONTEXT TYPE (TOPIC [dotted line] vs. NEUTRAL [solid line]) within each word order (SO vs. OS). Error bars represent the standard error of the mean.

Table 5.2: Statistical results of comprehensibility judgments (Exp. 1)
The table shows the statistical results for the fixed effects of the final logit mixed model analyses of the comprehensibility judgment data.

Fixed effects		*SE*	*z-value*	slope adjustments
Full factorial logit mixed model				
Intercept	-1.56	0.22	-6.97***	
CONTEXT TYPE (TOPIC vs. NEUTRAL)	0.60	0.20	3.05**	participants
WORD ORDER (SO vs. OS)	4.00	0.54	7.45***	participants, items
CONTEXT TYPE x WORD ORDER	0.53	0.20	2.64**	participants
Post hoc logit mixed models				
OS CONTEXT TYPE (TOPIC vs. NEUTRAL)	1.10	0.34	3.20***	participants
SO CONTEXT TYPE (TOPIC vs. NEUTRAL)	-0.20	0.18	-1.09	

Note: Significance levels: * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$. *b* = estimate, *SE* = standard error.

5.3 Experiment 2

In Experiment 2, participants were presented with the same stories as in Experiment 1, while ERPs were used to investigate the effect of the preceding discourse context (CONTEXT TYPE: topic vs. neutral) during online processing of German SO and OS sentences. Simultaneously, the behavioral performance of the participants was monitored in the form of a sentence-picture-verification task administered in 20 % of the trials.

5.3.1 Materials and methods

5.3.1.1 Participants

Twenty-one German native speakers (13 female, *M* age 25 years, age range 19-30 years) participated after giving informed consent. None of the participants took part in Experiment 1. All participants were right-handed as assessed by a German version of the Edinburgh Handedness Inventory (Oldfield, 1971). All had normal or corrected-to-normal vision and did not report any neurological disorder. Participants were reimbursed or received course credits for participation. Two participants were excluded from the analysis due to response accuracy scores below 60 % in the sentence-picture-verification task (see 3.1.3 Procedure). Data analysis was thus based on the remaining 19 participants (11 female, *M* age 25 years, age range 19-30 years).

5.3.1.2 Material

Material for Experiment 2 was identical to Experiment 1. Additionally, 32 colored drawings depicting the scene of the preceding target sentence with correct (matching) or exchanged (mismatching) thematic roles (e.g., The owl paints the hedgehog. vs. The hedgehog paints the owl.) were created for the sentence-picture-verification task. For each of the four experimental conditions (NEUTRAL SO/OS, TOPIC SO/OS) the same number of matching/mismatching pictures was constructed.

5.3.1.3 Procedure

The procedure was identical to that of Experiment 1 except for the following three methodological adjustments: First, the participant was prepared for EEG recording prior to the experiment. Second, presentation of the target sentence was preceded and followed by a fixation cross for 500 ms in the center of the screen to reduce vertical eye movements of the participant. Third, instead of the behavioral judgment task on story comprehensibility, the participants performed a sentence-picture-verification task that followed the target sentence in 20 % of the trials: After offset of the fixation cross, which followed the target sentence, the matching/mismatching picture was presented for two seconds before the participant had to press the corresponding button (yes vs. no) within a time window of two seconds. The assignment of the response buttons to the right index and middle fingers was counterbalanced across participants. A written instruction informed participants to read each scene attentively and silently and to answer the sentence-picture-verification task as accurately and fast as possible. Participants were asked to sit in a relaxed manner and to avoid blinks as well as other movements during sentence reading. The whole experimental session including three practice trials and pauses after each of the 40 trials lasted approximately 30 minutes plus electrode preparation.

5.3.1.4 EEG Recording

The EEG was recorded through a 32 channel active electrode system (Brain Products, Gilching, Germany) fixed at the scalp by means of a soft cap (Easycap, Inning, Germany). The electrode configuration included the following 29 scalp sites according to the

international 10-20 system (American Electroencephalographic Society, 2006): F7/8, F5/6, F3/4, FC3/4, C5/6, C3/4, CP5/6, P3/4, P7/8, PO3/4, FPz, AFz, Fz, FCz, Cz, CPz, Pz, POz, Oz. To detect blinks and vertical eye movements, an electrooculogram (EOG) was monitored by one electrode under and one electrode above the right eye. The ground electrode was placed at FP1. EEG data were acquired with a sampling rate of 1000 Hz. Impedances were kept below 5 kOhm. The left mastoid served as the reference electrode online, but the recording was re-referenced to bilateral mastoids offline.

5.3.1.5 ERP data analysis

For ERP data analysis, Brain Vision Analyzer software (version 2.0.2; Brain Products, Gilching, Germany) was used. EEG raw data were filtered by applying the Butterworth zero phase filter (low cutoff: 0.3 Hz; high cutoff: 70 Hz; slope: 12 dB/oct) to exclude slow signal drifts and muscle artifacts, and a notch filter of 50 Hz. Artifacts caused by vertical eye movements were corrected by the algorithm of Gratton, Coles and Donchin (1983). An automatic artifact rejection was used to reject blinks and drifts in the time window of -200 to 1500 ms relative to the onset of the critical stimuli in the target sentence: first determiner phrase (DP1), verb (V) and second determiner phrase (DP2) (rejection criteria: max. voltage step of 30 µV/ms, max. 200 µV difference of values in interval, lowest activity of 0.5 µV in intervals). Relative to the onset of DP1, V, and DP2, on average 5.71 % of trials were rejected with an equal distribution across onsets of critical stimuli and experimental conditions [$F(2, 36)$, $p > .1$] . ERPs were averaged for each participant and each condition within a 1500 ms time window time-locked to the onset of the critical stimuli with a 200 ms pre-stimulus onset baseline.

Based on visual inspection of the ERPs and according to the literature on language-related ERP components (i.e., P200, N400, Late Positivity), mean amplitude values of the ERPs per condition were statistically analyzed in the time windows 100-300 ms (P200), 300-500 ms (N400) and 500-700 ms (Late Positivity). The following nine regions of interest (ROIs) were computed via mean amplitudes of the three corresponding electrodes: left frontal (F7, F5, F3), left fronto-central (FC3, C5, C3), left centro-parietal (CP5, P3, P7), right frontal (F8, F6, F4), right fronto-central (FC4, C6, C4), right centro-parietal (CP6, P4, P8), frontal-midline (FPz, AFz, Fz), central midline

(FCz, Cz, CPz), parietal midline (Pz, POz, Oz). The statistical ERP analysis followed a hierarchical schema (e.g., Bornkessel et al., 2003; Rossi et al., 2011) using IBM SPSS Statistics (version 21.0). Firstly, a fully crossed repeated measures analysis of variance (ANOVA) with the factors CONTEXT TYPE (TOPIC, NEUTRAL), WORD ORDER (SO, OS), and ROI (nine levels) was computed separately for the three time windows post onset DP1, V, and DP2. We applied the correction of Greenhouse and Geisser (1959) and report the corrected F- and p-values but with the original degrees of freedom. Only statistically significant ($p \leq .05$) and marginally significant ($p \leq .06$) main effects and interactions including the factors CONTEXT TYPE and/or WORD ORDER were resolved in post hoc comparisons. Significant three-way interactions were resolved by computing ANOVAs on the next level. Whenever the ANOVA revealed a significant interaction of CONTEXT TYPE or WORD ORDER with ROI, paired t-tests were calculated to report the topographical distribution of the effect. As our study is concerned with the effect of CONTEXT TYPE within each WORD ORDER, a significant interaction of both factors would be resolved by WORD ORDER. With this procedure, we ensure to compare ERPs of identical DPs with regard to morphosyntax and thematic role. For presentation purposes only, the grand average ERPs displayed in Figure 5.2 and 5.3 were 7 Hz low-pass filtered (Butterworth zero phase filter: high cutoff: 7 Hz; slope: 12 dB/oct).

5.3.1.6 Behavioral data analysis

For statistical data analysis of the sentence-picture-verification task, logit mixed models for analysis of the binary distributed response accuracy data (correct vs. incorrect answers) were calculated. This statistical analysis followed the same procedure as described in Experiment 1.

5.3.2 Results

5.3.2.1 ERP results

Figure 5.2 displays the grand average ERPs at selected electrode positions of the respective ROIs time-locked to the onset of DP1. For complete statistical details of the

ERP analysis at DP1 see Table 5.3. Figure 5.3 shows the grand average ERPs of one selected exemplary electrode time-locked to the onset of the verb and DP2, respectively.

ERP results time-locked to onset of the first determiner phrase (DP1)

For ERPs in the time window 100-300 ms post onset DP1, the ANOVA including the factors CONTEXT TYPE (TOPIC vs. NEUTRAL) and WORD ORDER (SO vs. OS) and ROI revealed a significant main effect of CONTEXT TYPE [$F(1, 18) = 5.48$, $p \leq .05$]: If DP1 was preceded by the topic context, the positivity around 200 ms was reduced (compared to the neutral context). The ANOVA in the 300-500 ms time window yielded neither any statistically significant main effects nor interactions [$p > .1$]. For the 500-700 ms time window, the ANOVA revealed a significant interaction of WORD ORDER x ROI [$F(8, 144) = 4.14$, $p \leq .01$] as well as WORD ORDER x CONTEXT TYPE x ROI [$F(8, 144) = 4.15$, $p \leq .05$].[11] Separate post hoc analyses to resolve the three-way interaction of WORD ORDER x CONTEXT TYPE x ROI by WORD ORDER revealed a significant interaction of CONTEXT TYPE x ROI in sentences with OS order [$F(8, 144) = 2.99$, $p \leq .05$] (see Figure 5.2, lower panel). Follow-up t-tests showed a significantly reduced positivity from 500-700 ms for OS sentences preceded by the topic context relative to the neutral context in the right-frontal and frontal-midline ROI [$t(18) = -2.53/-2.28$, $p \leq .05$]. For SO sentences, the post hoc ANOVA did not show any significant differences in the ERPs with regard to the factor CONTEXT TYPE [$p > .1$] (see Figure 5.2, upper panel).

[11] Similar results were found for the analysis of the time window 500-900 ms.

Table 5.3: Results of analysis of variance (ANOVAs) of the ERPs in Exp. 2
The table shows the ANOVA results of the ERPs for different time windows (TW) time locked to onset of the first determiner phrase (DP1).

	df	*F-values*		
		TW 100-300 (P200)	TW 300-500 (N400)	TW 500-700 (P600)
CONTEXT TYPE	1, 18	5.48*	0.02	0.73
CONTEXT TYPE x ROI	8, 144	1.77	0.78	0.34
WORD ORDER	1, 18	1.99	1.04	0.67
WORD ORDER x ROI	8, 144	0.66	0.98	4.14**
WORD ORDER x CONTEXT TYPE	1, 18	2.56	0.47	3.07
WORD ORDER x CONTEXT TYPE x ROI	8, 144	1.68	2.09	4.15*

Note. Greenhouse & Geisser (1959) corrected significance levels: * $p \leq .05$; ** $p \leq .01$. *df* = degrees of freedom.

ERP results time-locked to onset of the verb

The ERPs in the three different time windows 100-300 ms, 300-500 ms and 500-700 ms post verb onset neither revealed any statistically significant main effects nor interactions with regard to the factors CONTEXT TYPE, WORD ORDER and/or ROI [$p > .1$]. As suggested by one anonymous reviewer, we performed an additional ERP analysis without baseline correction to account for possible baseline correction effects during the course of sentence processing (see e.g., Friederici, Wang, Herrmann, Maess, and Oertel (2000) and Wolff et al. (2008) for a similar procedure). The results revealed a significant main effect of WORD ORDER in the 100-300 ms time window [$F(1, 18) = 5.89$, $p \leq .05$] (OS more positive than SO) and a significant interaction of WORD ORDER x ROI in the 300-500 ms time window [$F(8, 144) = 3.25$, $p \leq .05$]. The post hoc t-test analysis to resolve the WORD ORDER x ROI interaction in the 300-500 ms time window revealed an enhanced negativity for OS compared to SO sentences in the left central ROI [$t(18) = 2.64$, $p \leq .05$] (see Figure 5.3 (left panel) for the grand average ERPs time-locked to the onset of the verb at an example electrode of the left central ROI).

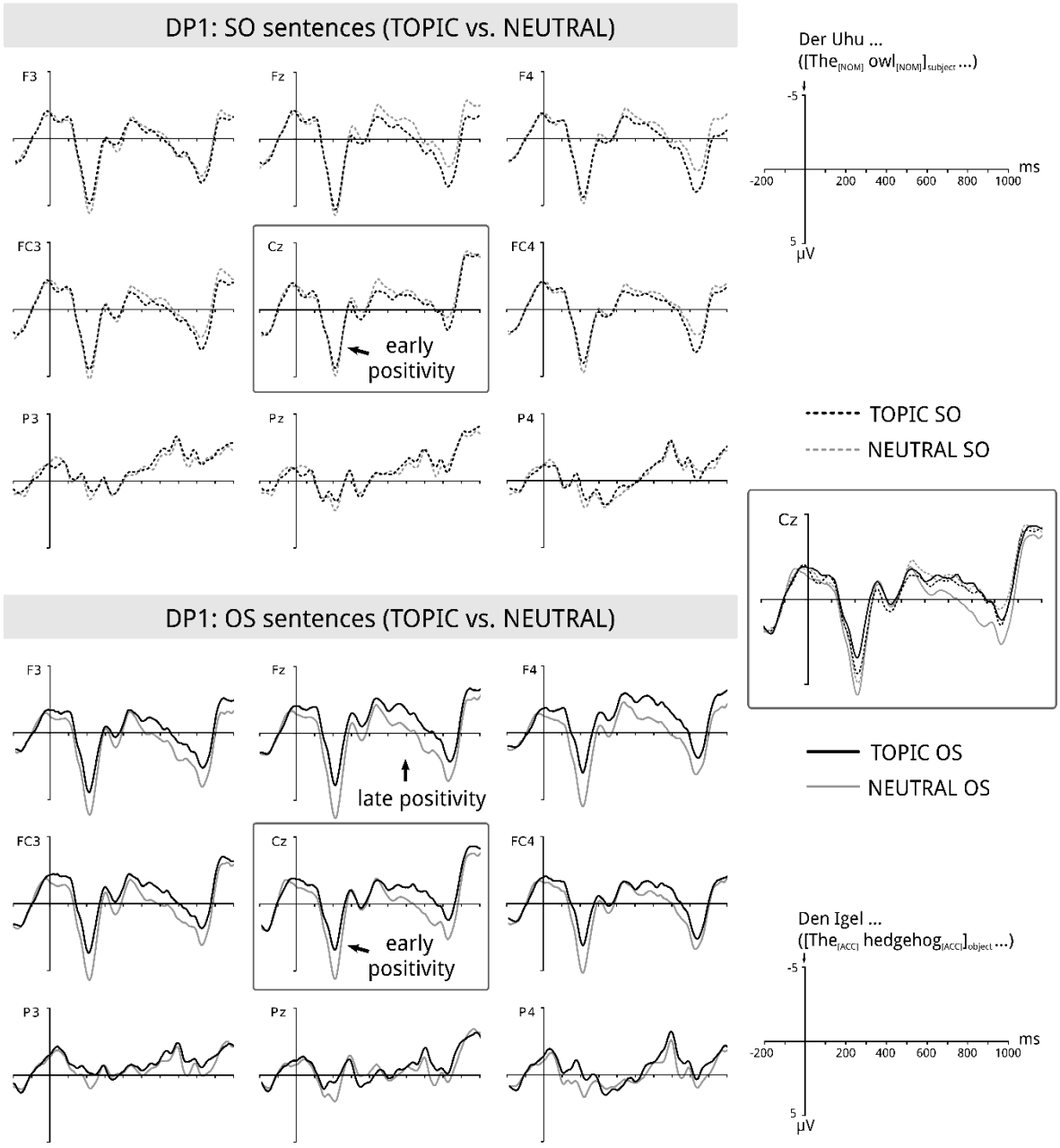

Figure 5.2: Grand average ERPs time-locked to DP1 (Exp. 2)
The Figure shows the grand average ERPs (with baseline correction) at selected electrodes time locked to the onset of the first determiner phrase (DP1) of the target sentence the effect of CONTEXT TYPE (TOPIC vs. NEUTRAL) within each word order (upper panel: TOPIC SO [dotted black] vs. NEUTRAL SO [dotted gray], lower panel: TOPIC OS [solid black] vs. NEUTRAL OS [solid gray]). Negativity is plotted upwards.

ERP results time-locked to onset of the second determiner phrase (DP2)

Statistical analysis of the ERPs time-locked to the onset of DP2 revealed a significant interaction of WORD ORDER x ROI in the time windows 300-500 ms [$F(8, 144) = 3.09$, $p \leq .05$] and 500-700 ms [$F(8, 144) = 3.53$, $p \leq .01$]. Post hoc t-tests showed that ERPs at DP2 were significantly more positive for OS sentences compared to SO sentences in the left frontal ROI for the 300-500 ms [$t(18) = -3.45$, $p \leq .01$] as well as for the 500-700 ms time window [$t(18) = -2.24$, $p \leq .05$].

Similar to the analysis with baseline correction, ERPs without baseline correction time-locked to the onset of DP2 showed the same pattern, but only in the later time window: The ANOVA of ERPs without baseline correction resulted in a marginally significant interaction of WORD ORDER x ROI [$F(8, 144) = 2.46$, $p \leq .06$] in the time window of 500-700 ms. As revealed by post hoc t-tests in this time window, the ERPs of OS sentences were significantly more positive compared to SO sentences in the frontal midline ROI [$t(18) = -2.12$, $p \leq .05$] (see right panel in Figure 5.3).

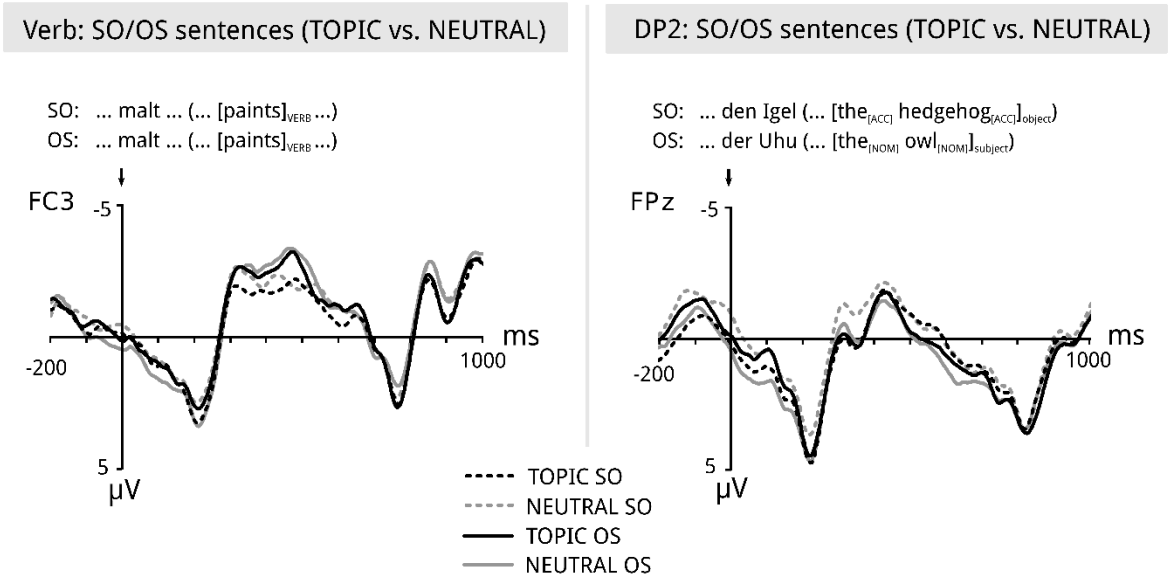

Figure 5.3: Grand average ERPs time-locked to verb and DP2 (Exp. 2)
Grand average ERPs (without baseline correction) are plotted at one selected electrode time-locked to the onset of verb (left panel) and second determiner phrase (DP2) (right panel) of the target sentence, respectively (TOPIC SO [dotted black] vs. NEUTRAL SO [dotted gray] vs. TOPIC OS [solid black] vs. NEUTRAL OS [solid gray]). Negativity is plotted upwards.

5.3.3 Behavioral results

Participants showed the following response accuracy for each condition (in 20 % of the trials): NEUTRAL SO: $M = 0.92$ ($SE = 0.02$), TOPIC SO: $M = 0.86$ ($SE = 0.02$), NEUTRAL OS: $M = 0.84$ ($SE = 0.03$), TOPIC OS: $M = 0.88$ ($SE = 0.02$). The final logit mixed model analysis of the raw response accuracy data including by-participant and by-item random intercepts did not reveal any statistically significant differences concerning the fixed effects CONTEXT TYPE ($b = 0.03$, $SE = 0.65$, $z = 0.05$, $p > .1$), WORD ORDER ($b = 0.84$, $SE = 0.65$, $z = 1.28$, $p > .1$), or the interaction CONTEXT TYPE x WORD ORDER ($b = 0.29$, $SE = 0.65$, $z = 0.45$, $p > .1$).

5.4 Discussion

In the present study, we used an offline comprehensibility judgment task (Experiment 1) to determine if discourse context affects the judgments concerning the overall comprehension of stories with German SO and OS sentences, and applied ERPs (Experiment 2) to characterize the time course of context-induced effects during online sentence comprehension. The discourse contexts depicted two characters in a fictitious scene and a context question related to either the topic status of one character (topic context) or a wide focus of the entire scene (neutral context).

In summary, the results of both experiments clearly revealed a statistically significant interaction of the factors CONTEXT TYPE and WORD ORDER. The results of the comprehensibility judgment task (Experiment 1) demonstrate the participants` judgments on the comprehensibility of stories with OS target sentences were significantly improved if presented together with the topic context as compared to the neutral context. As predicted, no context effects were evident for the comprehensibility judgments of stories with SO target sentences. In line with the judgment data, during online comprehension of OS target sentences, ERPs (Experiment 2) were significantly modulated by the previous topic context: Compared to neutral context, the topic context elicited a less pronounced Late Positivity at the sentence-initial object position (DP1). Thus, for the OS sentences, the processing of identical sentence structures was significantly affected by the preceding context type. As expected, no effect of context

was found during online processing of SO sentences; supporting the assumption that context information does not play a crucial role for processing of canonical word order. In addition, we observed a significant modulation of an early positivity peaking around 200 ms: Independent of word order, the early positive peak was reduced for target sentences following the topic relative to the neutral context. We interpret this finding as a perceptual mismatch response to repeated words (see below). Notably, in ERPs, the impact of context information during sentence processing was exclusively observable at the sentence-initial position (DP1) and did not elicit any further differential effects as the sentence unfolds (i.e., verb, DP2, for which we only found word order effects). In the following, we will discuss our results first in light of ERP components, before turning in more detail to word order effects and the impact of aboutness topic on the processing of non-canonical sentences.

5.4.1 Late Positivity

ERP studies investigating discourse level processing attributed the Late Positivity to processing costs for updating the current discourse model (e.g., Burkhardt, 2006, 2007; Cowles, 2003; Hirotani & Schumacher, 2011; Hung & Schumacher, 2012; Kaan et al., 2007; Schumacher & Hung, 2012; Wang & Schumacher, 2013). If the previously established discourse representation has to be updated by the listener, an increased Late Positivity has been induced. We suggest that establishing aboutness topic status of one of the two given characters by means of the context question increased the activation of this character in the present discourse model. Parallel to those recent ERP studies on discourse level processing and in line with the SDM, we interpret the Late Positivity in our study as an index for differential discourse updating costs of the established discourse model depending on whether the aboutness topic of the upcoming sentence has been announced previously.

In our study design, the topic context induced the expectation that the topic will be announced at the first position of the target sentence because the sentence-initial position is preferably filled by topic in German main clauses (e.g., Büring, 1999). If the first position of the target sentence is an object (i.e., OS sentence), fewer costs for updating the discourse model are induced if the sentence was preceded by a topic

context as compared to a neutral context. Hung and Schumacher (2014) have observed that, for Mandarin Chinese at least, presenting a less prominent referent in topic position caused higher updating costs as reflected in a Late Positivity. While Hung and Schumacher manipulated prominence in terms of animacy, it could be argued for our study that the topic context increased the information structural prominence of one of the two previously given referents (both animate). Hence in OS, the prominent announcement of the topic referent led to reduced updating costs of the mental model as compared to the neutral context, in which both referents were equally prominent – rendering none of them plausible to be placed in the sentence-initial object position. If the first position of the target sentence is a subject (i.e., SO sentence), there are no differential discourse updating costs dependent on the preceding context. We might not see a comparable modulation of the Late Positivity at the sentence-initial position in SO sentences, as –due to the strong subject-first-preference in German (e.g., Hemforth, 1993)– the canonical word order is felicitous and hence easy to process even in the absence of context information (see Subsections 51.1 and 5.1.3). The well-established interpretation of the Late Positivity in terms of the P600 (also syntactic positive shift, SPS) as reflecting syntax specific processing costs for structural reanalysis (e.g., Hagoort, 1993; Osterhout & Holcomb, 1992) and repair mechanisms (e.g., Friederici, Steinhauer, Mecklinger, & Meyer, 1998) is not sustainable for the Late Positivity in our study. In particular, we found the Late positivity was elicited during processing of the very same non-canonical structures in which neither syntactic anomalies (i.e., ambiguity resolution) nor violations (e.g., of the phrase structure) were present. Thus, this Late Positivity is in fact modulated by the preceding discourse level information and indexes discourse updating costs in line with the assumption of the SDM. The interpretation of the Late Positivity in our study is also compatible with the assumptions of the eADM: In the third phase of sentence processing Late Positivities indicate the integration of core-external (e.g., discourse) information and have been linked to the P300 family (Bornkessel & Schlesewsky, 2006b). P300 (or P3) responses are positive deflections of the ERP induced around 300 ms after stimulus onset (Sutton, Braren, Zubin, & John, 1965) typically evoked by infrequent stimuli in oddball paradigms (see Nieuwenhuis, Aston-Jones, & Cohen, 2005 for a review). In general,

amplitude and latency of the component are considered to be influenced by (unconscious) expectancy[12], task relevance, novelty, contextual constraints, and motivational significance (see e.g., Nieuwenhuis et al., 2005). Of most interest to our study, the P300 has been assumed to be related to domain-general context-updating processes and to reflect the revision of a mental model or the "conditions of the environment" (Donchin and Coles, 1988, p. 367; but see Verleger, 1988 and the following commentaries). Our design strictly followed a simple pattern of lead-in–context-question–target-sentence, revealing all referents given in the lead-in. The reduced Late Positivity in response to the sentence-initial object following the topic context could index a reduced need for general context updating, because the listener is less "surprised" about the object if previously announced as the topic of the scene compared to the neutral context. Thus, in line with Cowles (2003) who also reported a contextually modulated Late Positivity (i.e., the *Late Positive Component* (LPC)) during sentence comprehension, the Late Positivity in our study could reflect context-updating processes in terms of the P300. Notably, a number of authors argue against the context-updating interpretation of the P300 in favor of a general reflection of simple attentional, evaluative, or memory mechanisms (for a review, see Nieuwenhuis et al., 2005). Hence, it remains a matter of debate if Late Positivities/P600 responses elicited by sentences really belong to the P300 family or whether they should be considered an independent component (e.g., Coulson, King, & Kutas, 1998; Roehm, Bornkessel-Schlesewsky, Rösler, & Schlesewsky, 2007; ; see Brouwer, Fitz, & Hoeks, 2012 for a related discussion of the P600 in response to semantic violations or illusions).

5.4.2 N400 component

The N400 has been described as another ERP component sensitive to discourse level information. It is thought to reflect processing costs for linking an entity to the current mental model (Burkhardt, 2006; Burkhardt & Roehm, 2007; Wang & Schumacher, 2013). The SDM assumes that discourse linking processes are driven by expectancy as

[12] Note, the P300 is not just simply evoked by unexpected stimuli but also present in response to highly expected words, depending on task demands and individual processing strategies (Roehm, Bornkessel-Schlesewsky, Rösler, & Schlesewsky, 2007).

indexed by a modulation of the N400 (see Subsections 5.1.2 and 5.1.3). In these studies, the degree of inferability, expectancy, or accessibility of an entity in the mental model modulated the N400: The N400 for previously given, expected, or repeated noun phrases was reduced because those entities were easier to link to the current discourse. Importantly, due to the preceding lead-in context in our study which was identical for the neutral and the topic condition, both characters of the scene were discourse-given (Prince, 1981). Hence, we controlled for any processing advantages caused by the well-established given-before-new ordering principle (Clark & Haviland, 1977), or different degrees of inferability of an entity that might coincide with context effects (i.e., focus on subject or object) during processing of sentences with varying word order from previous studies (e.g., Bornkessel et al., 2003; Meng et al., 1999). Thus, absence of an N400 modulation in our study might be due to the fact that both characters of the scene were previously mentioned in the lead-in context, and thus equally expected and accessible in the mental model. This is in line with Burkhardt and Roehm (2007), who argue that both entities within a coordinated noun phrase –in our experimental design the two animals in the lead-in (e.g., *the owl and the hedgehog*)– evoke the same representational status in terms of accessibility or saliency in the mental model. In the framework of the SDM, our design was effective in the modulation of costs for updating the current discourse model (Late Positivity, see above) but not for expectancy-based discourse linking processes (N400).

5.4.3 Early positivity

Notably, in the topic condition, the topic of the context-question (e.g., *What about the owl?*) was directly repeated at the sentence initial position of the target sentence (SO and OS sentences), whereas such a repetition was not present in the target sentence following the neutral context (e.g., *What exactly is going on?*). Accordingly, the context type in our study revealed a broadly distributed early positive peak time-locked to the onset of the target sentence independent of its word order. As the topic context induced a reduction of this early positivity relative to the neutral context, we suggest that this context effect might be confounded with basic processes of information encoding due to word repetition in one but not the other context. The early positivity we found

showed a similar peak and latency pattern as the positivity around 200 ms (c.f., P200) for which mixed results regarding its functional nature are reported in dependence on the experimental paradigm (e.g., Coulson, Federmeier, van Petten, & Kutas, 2005; Federmeier & Kutas, 2001; Friedrich & Kotz, 2007). As early modulations of ERPs, such as the P200, have commonly been associated with processes of basic information encoding (for visual stimuli see for instance Dunn et al., 1998; Evans & Federmeier, 2007; Luck & Hillyard, 1994), we propose an interpretation of the reduced early positivity for repeated words in the topic condition in terms of a word repetition effect. Note that so far contradictory results have been reported with regard to amplitude and latency of ERPs elicited by word repetition: On the one hand side, some studies did not find a reduced but instead an enhanced early positivity for repeated words (see e.g., van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991). However, in line with our data, a reduced early positivity for repeated words was found in word lists (e.g., Nagy & Rugg, 1989; Rugg, 1985). Most interestingly, Hung and Schumacher (2012) reported a similar discourse-related early P200-effect as our study (i.e., an enhanced P200 for novel-topic > topic-shift > topic-continuity; see also Hung & Schumacher, 2014). They interpreted the P200 –which was reduced for processing similar graphical forms– as an early perceptual mismatch response. This is in line with our interpretation of the present finding in terms of an early perceptual repetition effect in the topic condition.

5.4.4 Word order effects

Some ERP studies examining word order variation in German main clauses (i.e., prefield) without a preceding context demonstrated processing difficulties in terms of an enhanced LAN for OS compared to SO at the first DP (e.g., Matzke et al., 2002; Rösler et al., 1998), whereas other studies did not report such an effect of canonicity (e.g., Frisch et al., 2002; Knoeferle et al., 2007). For the German middlefield, robust processing difficulties in form of the scrambling negativity for OS vs. SO are reported even if preceded by context information (e.g., Bornkessel et al., 2003; Bornkessel & Schlesewsky, 2006a).

As mentioned above, we did not focus on the direct comparison of the two word orders for the following reasons: First, SO is the canonical and more frequent word

order in German; any differences could hence be confounded by those effects. Second, grammatical and thematic role coincided in our material. Thus, we would not only compare word order but also the order of thematic roles. Therefore, we prefer to interpret our context effects within each word order to assure we compare the same target sentences. However, the ERPs in our study indicate that word order immediately interacted with the preceding context during incremental sentence processing, as reflected by the Late Positivity at DP1 -the position that immediately followed the context question and revealed the crucial case marking of subject/object and the thematic role. Hence, it seems that similar to Schumacher and Hung (2012) no processing difficulties for OS vs. SO in terms of a negative deflection at the sentence-initial position of German main clauses was elicited –if embedded in a strong licensing context. At both subsequent sentence positions (i.e., verb, DP2) a significant word order effect was found. OS (vs. SO) sentences elicited an early positivity (100-300 ms) as well as a left central negativity 300-500 ms after the finite verb and a frontally distributed positivity 500-700 ms after the DP2. Similar word order effects on ERPs at subsequent sentence positions have been reported in other studies (e.g., a negativity around 350-550 ms relative to verb onset (Wolff et al., 2008); a positivity (400-700 ms) at DP2 (Fiebach, Schlesewsky, & Friederici, 2002). In line with these studies, we interpret the word order effects in our study as reflecting general processing costs for OS compared to SO sentences.

5.4.5 Aboutness topic and sentence comprehension

In line with recent studies using either offline (e.g., Meng et al., 1999; Weskott et al., 2011) or online methods (e.g., Bornkessel et al., 2003; Schumacher & Hung, 2012), our study shows that the type of context information crucially affected offline comprehensibility judgments and online sentence processing of non-canonical sentences in German. Unlike previous studies, we manipulated the topic status of our referents in terms of explicitly announcing the aboutness topic of the upcoming sentence rather than also manipulating givenness and/or focus. Taking into consideration the results of both experiments, we argue that the information structural concept aboutness topic serves as a felicitous context for the comprehension of OS

declarative sentences. The indication of the topic in our study did not coincide with animacy-based prominence of the characters (Tomlin, 1986) that could have led to any additional ordering preferences (e.g., Bornkessel-Schlesewsky & Schlesewsky, 2009bBornkessel-Schlesewsky & Schlesewsky, 2009b; Hung & Schumacher, 2014; Lenerz, 1977). In our study, grammatical and thematic role coincided (the grammatical subject was always the agent, the grammatical object was always the patient at both sentence positions); therefore, it is important to note that we interpret our context effects within each word order. Information-structurally, the topic –what the sentence is about– is preferably announced at the sentence-initial position (e.g., Büring, 1999; Reinhart, 1981). A recent study (Bornkessel-Schlesewsky & Schlesewsky, 2012) confirmed that in German aboutness-based information correlates with word order in the prefield, while prominence-based information affects word order in the middlefield. In line with these properties, we found that topic status seemed to affect information packaging in the prefield: If the sentence-initial object in OS has been established as topic by the preceding context the non-canonical word order was felicitous. This impact of topic was detectable in the offline judgments, as stories containing the OS target sentence were judged as harder to comprehend without a supportive context (i.e., neutral context). In line with this, we interpret the reduced Late Positivity during online processing of OS sentences following the topic context as reflecting reduced discourse updating costs compared to the neutral context.

The reduction of the Late Positivity is in line with reduced costs for updating the discourse representation in the listener as assumed by the SDM (Schumacher & Hung, 2012; Wang & Schumacher, 2013) as well as by the eADM (Bornkessel & Schlesewsky, 2006b). Hence, our findings are further evidence that currently processed information is directly interpreted and incrementally integrated in relation to a previously established discourse representation and support assumptions of recent sentence processing models (eADM, SDM, ISPH by Cowles, 2003). Although we can only speculate about the underlying brain structures, the “New dorsal-ventral stream model of sentence comprehension” of Bornkessel-Schlesewsky and Schlesewsky (2013) would assume the following left hemispheric brain regions to engage in our task: The dorsal processing stream is responsible for the time-dependent syntactic computations

and actor identification. The frontal cortex is engaged in top-down-control and conflict resolution (hence, the establishment and updating of word-order-expectations). Anterior lIFG has been shown to correlate with aboutness information (Bornkessel-Schlesewsky et al., 2012). Parietal brain regions are involved in linking single sentences to the previous discourse. However, these assumptions would need to be tested systematically in the future with experimental techniques other than ERPs and comprehensibility judgments.

In summary, the results of the offline comprehensibility judgments are directly reflected during online processing of the sentence-initial topic in these sentences. Offline measures, such as behavioral judgments, most likely coincide with metalinguistic awareness (Schütze & Sprouse, 2013). The additional online measure using ERPs emphasizes the impact of the topic information on the processing of non-canonical sentences in German. Thus, our ERP findings add explanatory information regarding the subsequent steps of sentence comprehension modulated by preceding discourse information. As processing of non-canonical sentences was crucially modulated by the preceding topic context, we argue that the processing of specific syntactic structures (e.g., with varying word order) is sensitive to discourse level information. Our data nicely fit to the SDM (see Schumacher & Hung, 2012 or Wang & Schumacher, 2013) which assumes two core processes of referential processing: (1) During discourse linking the expectation of the listener immediately modulates the processing of incoming information to connect current information to previously given information (not modulated in our study). (2) During discourse updating, the listener updates the previously established internal discourse representation and adapts the syntax-discourse mapping accordingly. The aboutness topic in the present study effectively reduced the discourse updating costs as reflected in the Late Positivity in the non-canonical sentences and the higher comprehensibility judgments, even though all referents were given in the previous context.

5.5 Conclusion

The present study characterized the nature and time course of an aboutness topic context on the comprehension of German declarative sentences within fictitious discourses. For non-canonical, but not for canonical sentences, we found an impact of the topic context which indicated one of two previously given characters of the scene as the aboutness topic compared to a context in which a wide scope of the scene was induced (neutral context). The results of both experiments, the offline comprehensibility judgment task and the ERPs during online sentence processing, indicate that the topic context selectively facilitated comprehension of the non-canonical word order. In the ERPs, easier comprehension of OS sentences preceded by the topic context was detectable in terms of a reduced Late Positivity at the sentence-initial object position. This reduced Late Positivity is interpreted as reflecting less effortful processing demands for updating the current discourse model in case the aboutness topic entity has previously been integrated therein. The present study supports recent evidence that during online sentence processing listeners immediately take incoming discourse information into account and dynamically adapt their internal discourse representation.

5.6 Acknowledgements

This research was supported by the Collaborative Research Centre (SFB 632) `Information structure´ funded by the German Research Foundation (DFG). IW was supported by the Stifterverband für die Deutsche Wissenschaft (Claussen-Simon-Stiftung). We thank Franziska Machens and Tobias Busch for assistance in data acquisition and analysis as well as Dr. Christina Harzman for proof reading.

Chapter 6.

Study 2:

Visual attention-capture cue in depicted scenes fails to modulate online sentence processing[13]

Dialogue and Discourse, 10(2), (2019), 79–104.

JULIANE BURMESTER[1], KATHARINA SPALEK[2], AND ISABELL WARTENBURGER[3]

[1] Empirical childhood research, University of Potsdam, Potsdam, Germany
[2] Department of German Language and Linguistics, Humboldt-Universität zu Berlin, Berlin, Germany
[3] Cognitive Sciences, Department Linguistics, University of Potsdam, Potsdam, Germany

Abstract

Everyday communication is enriched by the visual environment that listeners concomitantly link to the linguistic input. If and when visual cues are integrated into the mental meaning representation of the communicative setting, is still unclear. In our

[13] This chapter is adopted from: Burmester, J., Spalek, K., and Wartenburger, I. (2019). Visual attention-capture cue in depicted scenes fails to modulate online sentence processing. *Dialogue and Discourse 10*(2): 79-104. DOI: 10.5087/dad. 2019.204.

earlier findings, the integration of linguistic cues (i.e., topic-hood of a discourse referent) reduced discourse updating costs of the mental representation as indicated by reduced sentence-initial processing costs of the non-canonical word order in German. In the present study we aimed to replicate our earlier findings by replacing the linguistic cue by a visual attention-capture cue that directs participants' attention to a depicted referent but is presented below the threshold of conscious perception. While this type of cue has previously been shown to modulate word order preferences in sentence production, we found no effects on sentence comprehension. We discuss possible theory-based reasons for the null effect of the implicit visual cue as well as methodological caveats and issues that should be considered in future research on multimodal meaning integration.

Keywords: linguistic vs. visual salience, accessibility, discourse processing, ERP

6.1 Introduction

Everyday communication is multimodal, comprising linguistic as well as extra-linguistic (e.g., visual) information. A growing branch of psycholinguistic research highlights effects of extra-linguistic cues (e.g., eye gaze or gestures) and the visual environment on language processing (e.g., Crocker et al., 2010; Nappa & Arnold, 2014; Spevack, Falandays, Batzloff, & Spivey, 2018; Staudte et al., 2014). By contrast, traditional models of sentence comprehension do not explicitly account for the role of visual attention during the comprehension process (e.g., Bornkessel & Schlesewsky, 2006; Friederici, 2002; Marslen-Wilson & Tyler, 1980). However, discourse models (or situational/mental models) go beyond sentence-level processing. Discourse models propose that during communication, interlocutors build a non-linguistic mental representation of relevant discourse referents and events based on the incoming linguistic and visual perceptual input amongst multiple other factors (e.g., Bower & Morrow, 1990; Gernsbacher, 1991; Grosz & Sidner, 1986; Johnson-Laird, 1980; Van Dijk & Kintsch, 1983; Zwaan, 2004; Zwaan & Radvansky, 1998). Therein, referents of high attentional state are assumed to be mentally represented with a higher degree of mental accessibility and/or a higher activation level compared to referents of low

attentional state (e.g., Arnold, 2010; Arnold & Lao, 2015; Givón, 1988; Gundel et al., 1993). We will refer to those referents of high attentional state as being more salient. For the present study, we differentiate between *linguistic salience* which is verbally induced by, for instance, subject-hood or topic-hood of a referent vs. *visual salience* which is induced by, for instance, *exogenous* visual cues to a depicted referent. Exogenous visual cues initiate a reflexive attention shift of the addressee to the location of a stimulus. In our study we used the terms *implicit* vs. *explicit* visual cues to enable a more precise distinction of exogenous visual cues, which were used in previous studies (analogous to the distinction by Myachykov et al., 2012, p. 3). Implicit visual cues are presented below the threshold of perception (i.e., subconsciously). Explicit visual cues are presented above the threshold of perception (i.e., consciously) (for an overview about neuronal modulations by stimulus-driven (i.e., sensory cue-based) visual attention mechanisms, see Corbetta & Shulman, 2002). With the present study, we aim to test if visual salience induced by an implicit visual attention-capture cue to a depicted referent impacts online sentence-initial processing in a similar way as it has previously been shown for linguistic salience (Burmester, Spalek, & Wartenburger, 2014). Hence, we raise the underlying question if the accessibility degree of mentally represented discourse referents is affected by this type of implicit visual cue or if this is limited to linguistic cues.

In the linguistic domain, information structure is used to make certain entities of the discourse more salient. For instance, topic or aboutness topic is an information structural concept describing the entity (e.g., a referent) the sentence is about, that is, topic is attributed to that part of information about which the speaker intends to increase the listener's knowledge (Gundel, 1985; Reinhart, 1981). Hence, topic is ascribed not solely a formal linguistic but also a cognitive concept that activates the listener's mental representation at the beginning of a sentence (Portner, 2007). In the majority of languages, salient information –in terms of the grammatical subject and/or topic of the sentence– dominantly occupies the sentence-initial position, because subjects and topics own a higher accessibility degree compared to their complements, that is, objects and comments (e.g., Bock & Warren, 1985; Dryer, 2013; Tomlin, 1995). German is a language with a strong subject-first preference (e.g., Hemforth, 1993;

Weber & Müller, 2004): The canonical word order in German main clauses is subject-verb-object (SO) (see example sentence (1)). Morphological case marking at the respective noun phrases enables the identification of the grammatical function of subject (via nominative case (NOM)) and object (via accusative case (ACC)) for masculine nouns. (Note that in the example sentences (1) and (2), the nouns *"Wal"* [whale] and *"Hai"* [shark] lack overt case affixes, while the determiners are overtly case marked, which nevertheless allows the unequivocal identification of subject and object.)

(1) SO: Der Wal streichelt den Hai.
[the$_{[NOM]}$ whale$_{[NOM]}$]$_{subject}$ [strokes]$_{verb}$ [the$_{[ACC]}$ shark$_{[ACC]}$]$_{object}$.
The whale strokes the shark.

(2) OS: Den Hai streichelt der Wal.
[the$_{[ACC]}$ shark$_{[ACC]}$]$_{object}$ [strokes]$_{verb}$ [the$_{[NOM]}$ whale$_{[NOM]}$]$_{subject}$.
The shark, the whale strokes.

Despite the strong subject-first preference in German, information structural characteristics allow reordering of sentential constituents such that the object can precede the subject (see example sentence (2) for a non-canonical object-verb-subject (OS) main clause). However, OS sentences in German are much less frequent than SO sentences (e.g., Bader & Häussler, 2010) and need a suitable context which increases the salience of the sentence-initial object. In our previous work, for instance, linguistic salience in short, fictitious stories of two animals was induced by a topic question (i.e., *"What about the shark?"*), which revealed one of two previously mentioned (i.e., discourse given) referents as the topic of the scene (Burmester et al., 2014). Compared to a neutral cue not indicating topic-hood but a wide focus (i.e., *"What exactly is going on?"*), subsequent online sentence-initial processing of OS sentences is eased. This facilitating impact of linguistic salience (i.e., topic-hood) is reflected in the event-related potentials (ERPs) in the form of a sentence-initial Late Positivity, which is attributed to reduced discourse updating costs (e.g., Schumacher & Hung, 2012).

In 2018 we directly compared linguistic and visual salience cues (Burmester, Sauermann, Spalek, & Wartenburger, 2018): Visual salience induced via an *explicit* gaze-shift of a virtual person to a depicted referent speeds sentence-initial reading times of German SO and OS sentences similar to linguistic salience induced via a topic cue. Hence, the sentence-initial processing ease was evident 1) independent of whether salience was induced linguistically or visually compared to a preceding neutral cue, and 2) independent of whether the salient referent is mentioned as the sentence-initial subject or object (Burmester et al., 2018). This is line with other studies supporting the view that utterance comprehension is facilitated when the speaker's gaze increases the visual salience of depicted referents (e.g., Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012; Staudte & Crocker, 2011). However, not only speakers' eye gaze, which provides explicit information about referential intentions (henceforth: intentional information), influences utterance comprehension, but also various other visual salience cues. Staudte et al. (2014) showed that listeners benefit from an explicit (non-gaze) arrow cue (henceforth: attentional information) during utterance comprehension similar to eye gaze. Both the arrow and the gaze cue effectively direct listeners' visual attention to a depicted object, to finally anticipate this salient object for an upcoming verbal reference (Staudte et al., 2014). Arnold and Lao (2015) showed that another abstract type of visual attentional cue (i.e., a black rectangle with a size of approximately 1.0° x 1.0° of visual angle[14] presented for 200 ms at the target referent's location) together with the position of the referent in the visual display manipulates listeners' trial-initial attention in depicted scenes. Still, when listeners interpret a subsequent pronoun, their trial-initial attention only secondarily influences which antecedent they select as the most accessible referent in discourse. Instead, pronoun interpretation is primarily driven by the linguistic cue of sentence-initial mention. Overall, such studies provide

[14] Note that in order to establish comparability of visual cues published in previous research, we calculated the visual angle of the cues used by Arnold and Lao (2015) and Myachykov et al. (2012) post hoc as these studies did not report the visual angle. Calculations of the visual angle account for the visual cue's size and distance from participants' eyes. For instance, for Arnold and Lao (2015) the calculation was based on a screen distance of 650 mm (22 - 34 inches reported), screen size width of 390 mm, screen resolution width of 1280 pixels, and cue size of 38 pixels:
$Visual\ angle\ = degrees(arctan\ (((38/2)/(650 * 1280/390)))) * 2\ = 1.0°$

evidence that *explicit* visual attentional cues affect sentence- and discourse-level processes, although to a different extend than linguistic cues.

Evidence in favour of the impact of *implicit* visual cues comes from language production studies. Here, implicit similar to explicit visual cues effectively manipulate speakers' attention to referents in a depicted scene. This manipulation of the speakers' attention is reflected in the sentential structure they choose in picture descriptions: Gleitman et al. (2007) used an implicit visual cue by means of a black rectangle with a size of approximately 0.5° x 0.5° of visual angle presented for about 60 - 75 ms at the location of one of two subsequently depicted referents. Other sentence production studies used explicit visual cues by means of a black arrow (Tomlin, 1995), red dot or referent preview (e.g., Myachykov et al., 2012; Turner & Rommetveit, 1968) followed by the presentation of referents that are performing a simple transitive action. As a result, these implicitly or explicitly cued referents are more salient or accessible[15] than other, uncued referents as reflected in a greater likelihood of salient referents being mentioned sentence-initially as the grammatical subject and/or topic of the sentence (e.g., Arnold, 1998, 2010; Tomlin, 1997). Both cue types even lead to production of otherwise disfavoured linguistic structures (in English). For instance, in cases where the patient of the transitive action is cued, speakers produce the less frequent passive voice with salient referents (i.e., the patient) in sentence-initial position.

While production and/or eye-tracking data indicate shifts in the addressee's attention, ERPs allow us to investigate whether and when during the course of sentence processing increased effort is needed. Numerous ERP studies have provided insights into underlying discourse-level mechanisms elicited by different types of linguistic cues during online sentence processing (e.g., Bornkessel et al., 2003; Burkhardt, 2006;

[15] Note that in contrast to *mental* accessibility (e.g., Ariel, 1988) which relates to discourse-level processing, these sentence production studies attribute the term accessibility to the lemma/conceptual level which corresponds to the retrieval of a referent's mental representation from memory or "representing potential referents in thought" (Bock & Warren, 1985, p. 47). Accordingly, the lemma of a more accessible referent is earlier retrieved from memory and hence mapped to a more prominent syntactic role (Levelt, 1989). However, we use the term accessibility from both the side of discourse and from the side of sentence production, since for both sides referent-related features such as animacy, linguistic or visual prominence contribute to a referent's high accessibility degree (e.g., Ariel, 1988; Prat-Sala & Branigan, 2000). Disentangling different concepts of accessibility is not within the scope of the present study.

Burmester et al., 2014; Kaan et al., 2007). Based on specific neural correlates, the *Syntax-Discourse Model* (Schumacher & Hung, 2012), as an instance of a neurocognitive account of discourse processing, specifies two temporally distinct processing mechanisms of meaning computation: discourse linking (N400) and discourse updating (Late Positivity). In Burmester et al. (2014), the facilitative impact of the linguistic salience cue (i.e., topic-hood) elicits a reduced Late Positivity around 500 - 700 ms time-locked to the sentence-initial position of OS sentences, but not of SO sentences. In line with the assumptions of the Syntax-Discourse Model, the reduced Late Positivity in the non-canonical OS word order is attributed to reduced processing costs for updating the current discourse model following the linguistic topic cue compared to the neutral cue. This interpretation of the Late Positivity as an index for integration and updating processes of mental representations is further supported by recent studies (e.g., Delogu et al., 2018; or within the *neurocomputational model* of language comprehension by Brouwer et al., 2017). However, the assumptions of the Syntax-Discourse Model as well as of other discourse models (Hagoort & Van Berkum, 2007) go beyond the impact of purely sentential context on meaning computation, but include situational context information. Even more explicitly, the *Coordinated Interplay Account* (Crocker et al., 2010) highlights the role of visual attention for listeners' mental representations. This account assumes closely temporally synchronised stages of visual and linguistic information processing during sentence comprehension as supported by multiple "visual world" eye-tracking studies (e.g., Knoeferle & Kreysa, 2012) and also ERP studies (e.g., Knoeferle et al., 2007). For instance, *visual* cues reduced online processing costs of OS sentences: Facilitating cues included explicit, intentional, speech-aligned (beat) gesture cues indicating a specific sentence part as salient (Holle et al., 2012), or explicit visual presentations of the visually depicted event of the target sentence (Knoeferle et al., 2007). To the best of our knowledge it has not been reported so far how implicit visual cues that purely direct the addressee's attention to depicted referents impact online sentence processing.

Using ERPs for investigating the impact of implicit visual cues on sentence comprehension might contribute to our understanding of the underlying neurophysiological mechanisms during sentence processing which might be

comparable to those evoked by linguistic cues. Our study aims to answer the question whether –parallel to our earlier findings concerning linguistic salience– a referent in sentence-initial (i.e., topic) position is easier to process if visual salience is induced via an implicit attention-capture cue. Hence, by using an implicit visual cue in the present study we intend to conceptually replicate our earlier ERP-findings, that is, the sentence-initial Late Positivity modulation evoked by an (explicit) linguistic cue (Burmester et al., 2014). The implicit visual cue of the current study was presented for 66 ms analogously to the Gleitman et al. (2007) study in which a similar type of cue significantly manipulated speakers' attention in depicted scenes, and hence, modulated what speakers mentioned first during sentence production. In accordance with the earlier findings concerning linguistic salience, we predict that visual salience of a depicted referent induces modulations of the Late Positivity at sentence-initial position of subsequent OS sentences. Besides the Late Positivity, the linguistic topic cue in Burmester et al. (2014) elicited an early perceptual repetition effect due to word repetition in the topic but not in the neutral condition. This effect was reflected in a reduced early positivity around 200 ms at sentence-initial position of both SO and OS sentences. In the present visual cueing paradigm, no word repetition occurs. Therefore, we do not expect any modulations of this early positivity. In addition, the Burmester et al. (2014) study revealed a word order effect in terms of generally greater processing costs for OS than SO sentences that we expect to replicate in the present study.

6.2 Materials and methods

6.2.1 Participants

Thirty-one native speakers of German participated after giving informed consent. Except for one participant, participants were right-handed as assessed by a German version of the Edinburgh Handedness Inventory (Oldfield, 1971). All had normal or corrected-to-normal vision and had no reported neurological disorder. Participants were reimbursed or received course credits for participation. Data of two participants were excluded from further analysis, that is, one participant due to left handedness, and one participant due to a technical error during recording the electroencephalogram

(EEG). The analysed group consisted of 29 participants (15 female, mean age 24.8 years, age range 19.4 - 25.2 years).

6.2.2 Design and material

In the present study (analogous to Burmester et al., 2014) participants were presented with short stories of two animals that were going to perform a fictitious transitive action (e.g., a whale and a shark, one of which is going to stroke the other) while an EEG was recorded to investigate ERPs during online sentence processing. In contrast to Burmester et al. (2014), stories were additionally depicted by pictures of the two animals and the action instrument (cf. Figure 6.1). The study used a 2 x 2 within-subject design with the fully crossed factors CUE (TOPIC vs. NEUTRAL) and WORD ORDER (SO vs. OS sentences), resulting in four conditions: TOPIC SO, NEUTRAL SO, TOPIC OS, NEUTRAL OS. A total of 160 different stories (40 per condition) was created based on coloured pictures of 40 animals (monomorphemic nouns of masculine gender which were 1-syllabic (*n* = 18) or 2-syllabic (*n* = 22)) and 10 actions (which were monomorphemic, 2-syllabic transitive, and accusative-assigning verbs). For 90% of nouns, NOM and ACC case were overtly marked only at the determiner. In the remaining 10 % of nouns, NOM case was overtly marked only at the determiner, but ACC case was overtly marked at the determiner *and* noun (e.g., *"den Löwen"* [the$_{[ACC]}$ lion$_{[ACC]}$]$_{object}$). Nouns and verbs were controlled for normalised written lemma and type frequency values according to the dlex database (Heister et al., 2011). Moreover, other semantic and discourse factors such as animacy and discourse-givenness of sentential arguments impact referent accessibility and hence the ordering principles at the sentence-level (e.g., Clark & Clark, 1977; Grewe et al., 2006). We controlled for these factors by exclusively choosing animate referents that were explicitly mentioned in the lead-in sentence.

Each trial started with a red fixation cross signalling the beginning of a new story. Afterwards a blank screen for 500 ms was followed by a phrase-wise presented lead-in sentence (see Figure 6.1 (1)) introducing the two relevant animals of the scene, the action instrument and a corresponding prepositional phrase (e.g., the place where the animals were finding the action instrument). With regard to information structure, the lead-in revealed both animals as discourse-given (Prince, 1981) and the action as

inferable based on the mentioned instrument (Prince, 1992). The following visual context (2) consisted of the implicit visual attention-capture cue located at one of three picture positions: (i) the upper left or (ii) upper right animal (i.e., TOPIC CUE, respectively), (iii) the bottom centre position of the action instrument (i.e., NEUTRAL CUE). The visual cue was presented in the form of a black square with approximately 0.3° x 0.3° of visual angle presented against a light greyish background colour for a duration of 66 ms. The cue was immediately followed by the pictures at the respective positions similar to previous visual cueing paradigms (Gleitman et al., 2007; Myachykov et al., 2012). Hence, either one of the two animals was cued in order to direct participants' attention to the topic referent (TOPIC CUE), or the action instrument was cued in order to direct participants' attention to a wider scope of the scene, the to-be-performed action of the two animals (NEUTRAL CUE). The coloured pictures of the animals and actions were presented with 9.3° x 9.3° of visual angle. After a black fixation cross, the target sentence (3) was presented phrase-wise either in SO or OS WORD ORDER describing the thematic role relations of the depicted animals (i.e., who is performing the action with whom), followed by a blank screen for 200 ms. The target sentence consisted of a first determiner phrase (DP1), verb, second determiner phrase (DP2), and a prepositional phrase specifying the animals' location or the action instrument. DP1 was either both subject and agent of the action or object and patient/undergoer of the action. DP2 always carried the inverse syntactic and thematic role of DP1. Note here that in our study syntactic and thematic role always coincided. With the closing prepositional phrase (e.g., *"with the brush"*) we aimed to prevent that processing of DP2 is contaminated by "wrap up" effects typically occurring at the end of sentences (e.g., Just & Carpenter, 1980). The phrase-wise presentation durations were chosen in analogy to previous studies (Bornkessel et al., 2003): DPs and prepositional phrases presented for 500 ms, respectively; conjunctions, auxiliary verbs, and main verbs presented for 450 ms, respectively; with a 100 ms interstimulus interval. In 20 % of trials a sentence-picture-verification task (4) probed participants' attentive reading of the stories. For this, 32 pictures (eight per condition) depicting the content of the preceding target sentence were created -half with correct, half with exchanged (i.e., incorrect) thematic role assignments (e.g., *shark stroking whale* vs. *whale stroking*

shark). Pictures of the sentence-picture-verification were presented for 2 s before participants had to press the corresponding button within a 2 s time window. The verification task was followed by a blank screen for 500 ms. The experimental items were identical to the ones in Burmester et al. (2014) except that the lead-in sentence was not presented in a self-paced-reading manner but automatically and that the linguistic context question was replaced by the presentation of the visual context.

Due to the fictive character of the stories as in children's books, both animals could be plausible agents or patients of the action. In the visual context, animals of a story were always facing each other. In the target sentence, animals equally often occurred as the agent or patient of an action. Animals were distributed equally across conditions and were always performing the action with a different animal. Introducing the animal first or second in the lead-in sentence as well as the presentation of the animals on the left or right side of the screen was counterbalanced across conditions. To avoid possible effects of structural priming (e.g., Scheepers & Crocker, 2004), trials were presented in pseudo-randomized order with maximally two consecutive trials of the same condition and word order in the target sentence. Preferences of thematic role assignment or topic continuity due to preceding trials were minimized by at least five intermediate trials before an animal was repeated. There were four lists of 160 trials each. Lists were created such that within each list each item (i.e., animal pair and action) occurred once and across the four lists each item occurred in each condition. We presented each participant with one of these lists of 160 trials. These lists did not include any filler trials to arrange the experimental session in an appropriate time frame for participants' motivation and concentration ability, and hence, minimize artefacts and alpha waves in the EEG signal.

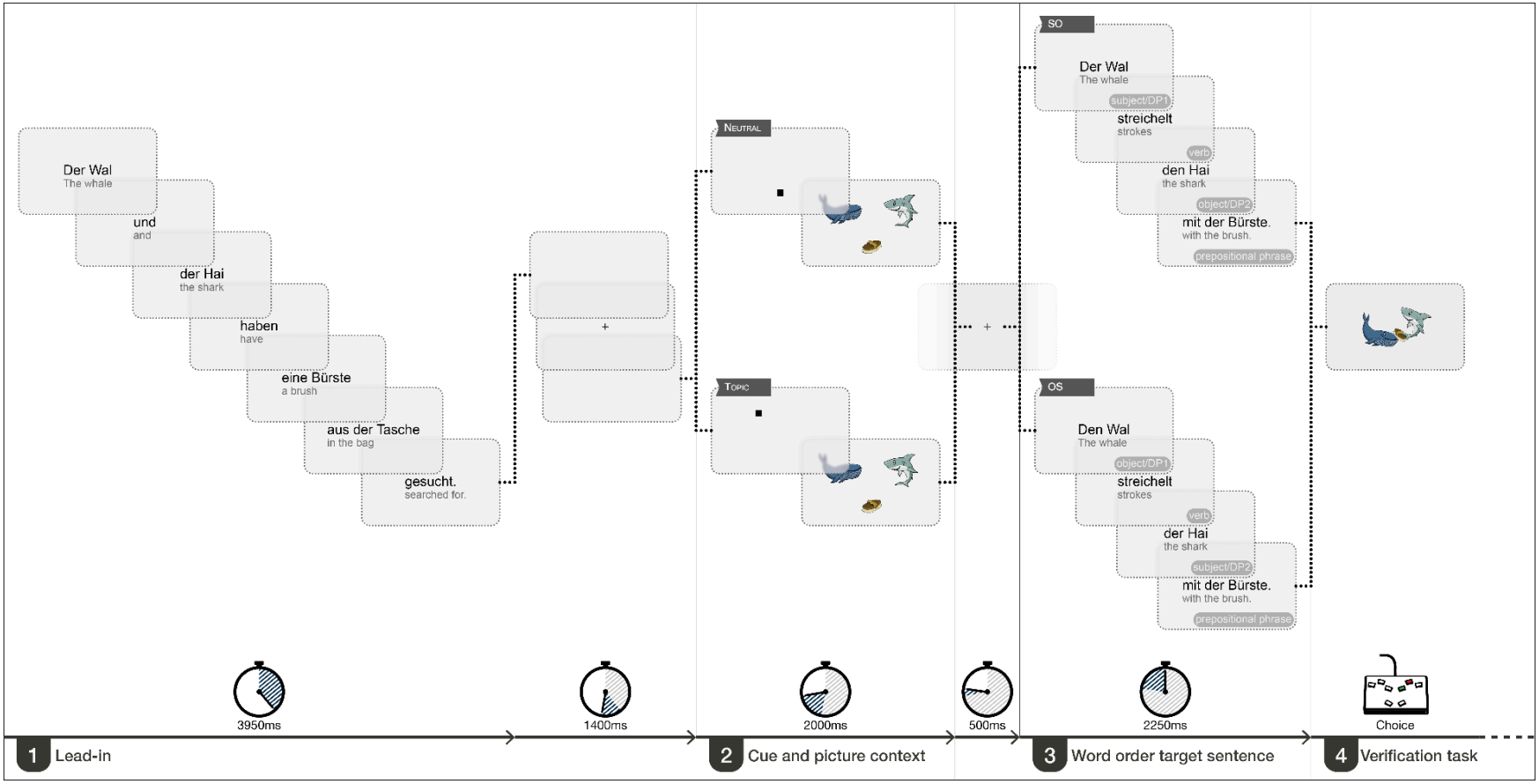

Figure 6.1: Experimental design of study 2 (implicit visual cue)
Experimental design of sample trial with approximate English translation written in grey. (1) Participants read the phrase-wise presented lead-in sentence followed by a blank screen (200 ms), fixation cross (1000 ms), and blank screen (200 ms). (2) The implicit visual cue presented for 66 ms (TOPIC vs. NEUTRAL CUE) was directly followed by the pictures. After another fixation cross, (3) the SO or OS WORD ORDER target sentence was presented phrase-wise. In the TOPIC cue condition, the SO or OS sentence mentioned the cued referent (i.e., whale) first (i.e., as DP1). (4) In 20 % of trials participants had to answer a sentence-picture-verification task afterwards.

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

6.2.3 Procedure

Participants were tested individually, seated in a sound-attenuated booth with 80 cm distance to a computer screen (1680 x 1050 pixels screen resolution). After the preparation for EEG recording, participants were visually presented (on screen) with all pictures used in the subsequent experiment with their corresponding word forms to become familiar with the pictures, that is, the 40 animals and 10 actions. Afterwards participants received a written instruction in which they were asked to read each story attentively and silently and to answer the sentence-picture-verification task after some of the stories as accurately and fast as possible. Participants were asked to sit relaxed, and to avoid eye-movements, blinks, and other muscle movements. Participants had a button box (Cedrus® response pad model RB-830) on their lap and performed three practice trials to become familiar with the procedure. To answer the sentence-picture-verification task, the green and red response button (according to correct vs. incorrect pictures) were assigned to the right fore and middle finger (which was counterbalanced across participants). Participants were instructed that they will be presented with a new story as soon as they see a red fixation cross and they press the yellow button of the button box on which they should put their left thumb throughout the whole experiment. The experiment was visually presented by means of the Presentation® software (version 14.1; www.neurobs.com). The whole experiment included pauses after each 40 trials and lasted approximately 30 minutes. In a post-experiment questionnaire, participants were asked if they have an idea about the purpose of the study and if they noticed anything in the course of the experiment, for instance, any cues or disturbances during picture presentation.

6.2.4 EEG recording

The EEG was recorded using a 32 channel active electrode system (Brain Products, Gilching, Germany) with a sampling rate of 1000 Hz. The electrodes fixed at scalp by means of a soft cap included the following 29 scalp sites according to the international 10-20 system (American Electroencephalographic Society, 2006): F7/8, F5/6, F3/4, FC3/4, C5/6, C3/4, CP5/6, P3/4, P7/8, PO3/4, FPz, AFz, Fz, FCz, Cz, CPz, Pz, POz, Oz. The electrooculogram (EOG) was monitored by electrodes above (position FP2) and below

the right eye. The ground electrode was placed at FP1. Impedances were kept below 5 kOhm. The left mastoid served as the reference electrode online, whereas recording was re-referenced to bilateral mastoids offline.

6.2.5 ERP data analysis

For ERP data analysis, the Brain Vision Analyzer software (version 2.1, Brain Products Gilching, Germany) was used. To exclude slow signal drifts and muscle artifacts from the EEG raw data the Butterworth zero phase filter (low cutoff: 0.3 Hz; high cutoff: 70 Hz; slope: 12 dB/oct) was applied additional to the notch filter of 50 Hz. For the correction of artifacts caused by vertical eye movements the algorithm by Gratton et al. (1983) was applied. We applied an automatic artifact rejection to reject blinks and drifts in the time window of -200 to 2150 ms relative to the onset of the target sentence as well as -200 to 500 ms relative to onset of the visual cue (rejection criteria: max. voltage step of 30 µV/ms, max. 200 µV difference of values in intervals, lowest activity of 0.5 µV in intervals). On average 1.64 % of trials was rejected. ERPs were averaged for each participant and each condition within a 2150 ms time window time-locked to the onset of the target sentence and within a 500 ms time window time-locked to the onset of the visual cue, with a 200 ms pre-stimulus onset baseline, respectively.

For the statistical analysis, IBM SPSS Statistics (version 25.0) was used. The chosen parameters for ERP analysis were identical to the ones used by Burmester et al. (2014) to maintain comparability with the ERP results on the impact of linguistic context information on online sentence processing, in which the same sentences were presented to participants. Hence, based on previous psycholinguistic research, we analysed language-related ERP components of the target sentence in the following time windows time-locked to the onset of DP1, verb, and DP2, respectively: 100 - 300 ms (P200), 300 - 500 ms (N400), 500 - 700 ms (Late Positivity). Via computation of mean amplitudes of three electrodes, respectively, nine regions of interest (ROIs, which were identical to the ones in Burmester et al., 2014) entered the statistical ERP analysis as the fixed factor ROI: LEFT FRONTAL (F7, F5, F3), LEFT FRONTO-CENTRAL (FC3, C5, C3), LEFT CENTRO-PARIETAL (CP5, P3, P7), RIGHT FRONTAL (F8, F6, F4), RIGHT FRONTO-CENTRAL (FC4, C6, C4), RIGHT CENTRO-PARIETAL (CP6, P4, P8), FRONTAL-MIDLINE (FPz, AFz, Fz), CENTRAL MIDLINE

(FCz, Cz, CPz), PARIETAL MIDLINE (Pz, POz, Oz). For statistical ERP analysis, mean amplitude values of ERPs within each condition were analysed following a hierarchical schema (e.g., Bornkessel et al., 2003; Burmester et al., 2014). Firstly, we computed a fully crossed repeated measures analysis of variance (ANOVA) with the fixed factors CUE (TOPIC vs. NEUTRAL), WORD ORDER (SO, OS), and ROI (nine levels, see above) for each of the three time windows time-locked to the onset of DP1, verb, and DP2, respectively.

In addition to the analyses of the target sentence, we analysed early ERP components (i.e., N1, P2) relative to the onset of the visual cue, henceforth termed CUE POSITION (i.e., TOPIC LEFT, TOPIC RIGHT, and NEUTRAL BOTTOM), in the time window of 100 - 200 ms as well as of 250 - 350 ms (see e.g., Luck & Hillyard, 1994; Mangun, 1995 for attention-based early visual processing changes reflected in different early evoked potentials). Notably, the pictures followed the cue immediately, hence these time windows started 100 ms and 250 ms after cue onset but also 34 ms and 184 ms after picture onset.

We report Greenhouse and Geisser (1959) corrected *F*- and *p*-values, the original degrees of freedom *(df)* in brackets, and the Greenhouse and Geisser epsilon *(ε)* factor for non-sphericity adjustments of the original *df* according to Jennings and Wood (1976) (only for *F*-tests with more than one *df* in the numerator). Statistically significant effects (i.e., $p < .05$) involving an interaction with ROI were resolved by computing post hoc paired t-tests to reveal the topographical distribution of the effect. We controlled for the Type I error due to multiple pairwise *t*-tests of levels of the fixed effects in the nine ROIs by adjusting the significance level according to the Bonferroni correction. Thus, for post hoc *t*-tests the following Bonferroni adjusted *p*-values (two-tailed) were considered as statistically significant at $\alpha = .05$: $p < .006$ to resolve the WORD ORDER x ROI interaction, and $p < .002$ to resolve the CUE POSITION x ROI interaction. For presentation purposes the displayed ERPs in Figure 6.2, 6.3, and 6.4 in Appendix B are 10 Hz low-pass filtered.

6.2.6 Behavioural data analysis

For the statistical analysis of the response accuracy of the sentence-picture-verification task, logit mixed models fitted by the Laplace approximation were calculated using the

lme4 package (Bates, Mächler, Bolker, & Walker, 2015) provided by the *R* environment (version 3.5.1, R Core Team, 2013). To analyse the binary distributed response accuracy data (correct vs. incorrect) with the logit mixed models, CUE, WORD ORDER, and the interaction of both were defined as fixed effects, and Participants and Items were defined as random effects. Fixed effects were coded as +/−.5 to resemble the contrast coding of traditional ANOVA analyses. Model fitting started with the simple model (i.e., the two fixed effects and their interaction, and Participants and Items as random intercepts). In a step-wise manner, slope-adjustments were included if they significantly improved the explanatory power of the simpler model without that slope adjustment as revealed by log-likelihood tests (e.g., Baayen, 2008). The statistics of the fixed effects of the final models are reported with estimates *(b)*, standard errors *(SE)*, *z*- and *p*-values.

6.3 Results

As reported in the post-experiment questionnaire, participants did not notice any manipulation of their visual attention suggesting that the cue was not consciously perceived and hence, truly implicit. In this Section, we first describe the ERP results with respect to initial and subsequent processing of the target sentence, before reporting the results of an additional analysis of the present ERP data following the visual cue together with the published data following the linguistic cue (Burmester et al., 2014). Secondly, we present the ERP results with respect to the onset of the visual cue. Thirdly, we report the behavioural results of the probe sentence-picture-verification task.

6.3.1 ERP results of the target sentence

Figure 6.2 illustrates the grand average ERPs at one representative electrode time-locked to the onset of the target sentence (i.e., DP1) followed by the subsequent sentence positions (i.e., verb and DP2) for both CUES (TOPIC and NEUTRAL) and both WORD ORDERS (SO and OS). For grand average ERPs at selected electrodes of each ROI, see Figure 6.4 in Appendix B illustrating the CUES *within* each WORD ORDER.

6.3.2 ERP results of sentence processing following the implicit visual cue

For sentence-initial processing, statistical analyses in the time windows of 100 - 300 ms and 300 - 500 ms time-locked to the onset of DP1 neither revealed any statistically significant main effects of CUE (TOPIC vs. NEUTRAL) or WORD ORDER (SO vs. OS), nor significant interactions of CUE, WORD ORDER and/or ROI [$p > .1$] (see Appendix A for the complete statistical output). The analysis in the following time window of 500 - 700 ms revealed a statistically significant main effect of WORD ORDER [$F(1, 28) = 6.254$, $p = .019$] and a significant interaction of WORD ORDER x ROI [$F(8, 224) = 3.004$, $p = .029$, $\varepsilon = 0.424$], but no statistically significant effects or interactions of the factor CUE. Separate post hoc analyses for SO and OS sentences (averaged across the cue conditions) within each ROI yielded a statistically significant enhanced positive-going ERP for OS sentences compared to SO sentences in the LEFT CENTRAL ROI [$t(28) = -3.605$, $p = .001$] and in the MIDLINE CENTRAL ROI [$t(28) = -3.369$, $p = .002$].

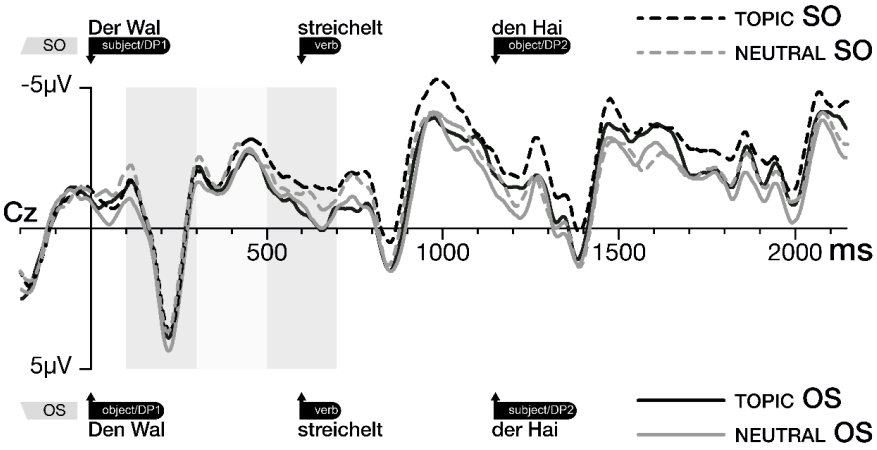

Figure 6.2: Grand average ERPs at electrode Cz time-locked to the onset of the target sentence
Grand average ERPs (with baseline correction) at one representative electrode of the MIDLINE CENTRAL ROI for the factors CUE (TOPIC vs. NEUTRAL) and WORD ORDER (SO vs. OS) time-locked to the onset of the target sentence, that is, the first Determiner phrase (DP1) followed by the verb and second Determiner phrase (DP2): TOPIC SO [dotted black] vs. NEUTRAL SO [dotted grey], TOPIC OS [solid black] vs. NEUTRAL OS [solid grey]. Negativity is plotted upwards. Grey shades indicate the three time windows of DP1.

For subsequent sentence positions, the statistical ERP analysis for the time windows 100 - 300 ms, 300 - 500 ms, and 500 - 700 ms post verb onset neither revealed a

statistically significant main effect nor interaction involving the factor CUE, but revealed statistically significant effects involving the factor WORD ORDER: In the time window 100 - 300 ms post verb onset a significant main effect of WORD ORDER [$F(1, 28) = 10.856$, $p = .003$] and a significant interaction of ROI x WORD ORDER [$F(8, 224) = 3.359$, $p = .013$, $\varepsilon = 0.485$] was reflected in an enhanced positive-going ERP for OS compared to SO sentences in the LEFT CENTRAL [$t(28) = -3.356$, $p = .002$] and LEFT POSTERIOR [$t(28) = -4.507$, $p < .001$] ROI, the RIGHT POSTERIOR ROI [$t(28) = -3.090$, $p = .004$], and the MIDLINE CENTRAL [$t(28) = -3.854$, $p = .001$] and MIDLINE POSTERIOR [$t(28) = -4.192$, $p < .001$] ROI. In the following time window of 300 - 500 ms post verb onset a significant interaction of ROI x WORD ORDER [$F(8, 224) = 8.101$, $p < .001$, $\varepsilon = 0.447$] was similarly reflected in form of a significantly enhanced positive ERP for OS compared to SO sentences in the LEFT POSTERIOR ROI [$t(28) = -3.448$, $p = .002$] and in the MIDLINE POSTERIOR ROI [$t(28) = -4.602$, $p < .001$]. Furthermore, in the time window of 500 - 700 ms post verb onset the significant interaction of ROI x WORD ORDER [$F(8, 224) = 6.157$, $p < .001$, $\varepsilon = 0.558$] was reflected in an enhanced positive ERP for OS compared to SO sentences in the MIDLINE POSTERIOR ROI [$t(28) = -3.114$, $p = .004$]. Statistical analyses of ERPs post onset of DP2 neither showed any statistically significant main effects, nor interactions of CUE or WORD ORDER in any of the calculated time windows (i.e., 100 - 300 ms, 300 - 500 ms, and 500 - 700 ms) [$p < .1$].

In summary, the ERP results of all three sentence positions (i.e., DP1, verb, DP2) did not show any statistically significant modulation by the preceding visual cue. An impact of the varying word order with enhanced positive-going ERPs for OS compared to SO sentences was evident in multiple time windows time-locked to the onset of DP1 (i.e., 500 - 700 ms) and verb (i.e., 100 - 300 ms, 300 - 500 ms, and 500 - 700 ms).

6.3.3 ERP results compared to sentence processing of the linguistic cue (Burmester et al., 2014)

Since the identical sentence material was used in the present study with the visual cue as in the study with the linguistic cue (Burmester et al., 2014), we aimed at directly comparing the impact of the visual vs. linguistic cue *modality* on sentence-initial processing. For this purpose, we computed additional comparisons of the published

ERP data following the linguistic cue with the ERP data following the visual cue by adding the between-subject factor MODALITY (VISUAL vs. LINGUISTIC) while the within-subject-factors CUE, WORD ORDER, and ROI maintained as in the original analysis (reported in Section 6.2.5). The results as summarised in Appendix C show the following statistically significant effects of MODALITY: In the 100 - 300 ms time window, the analysis showed a significant interaction of MODALITY x ROI [$F(1, 46) = 3.190$, $p = .038$] and of MODALITY x WORD ORDER [$F(1, 46) = 5.793$, $p = .020$], which was reflected in a marginally significantly different processing of SO vs. OS sentences following the linguistic cue [$t(18) = 2.092$, $p = .051$], which was absent following the visual cue [$t(28) = -4.16$, $p = .680$]. In line with the separate analyses, there were no statistically significant effects of MODALITY in the 300 - 500 ms time window. For the 500 - 700 ms time window post onset of DP1, the joint analysis revealed a marginally statistically significant interaction of MODALITY x CUE x WORD ORDER [$F(1, 46) = 3.963$, $p = .052$]. This confirms the results of the two separate analyses of the linguistic and visual cue: the presence of a statistically significant interaction of CUE x WORD ORDER x ROI [$F(8, 144) = 4.15$, $p < .05$] following the linguistic cue and its absence following the visual cue [$F(8, 224) = 1.605$, $p = 0.189$; $\varepsilon = 0.417$] (cf. Figure 6.2 and 6.4 in Appendix B) in the present study and Figure 2 and Table 3 in Burmester et al., 2014).

In summary, the visual cue had no impact on sentence processing in the present study. Therefore, the visual cue was not effectively increasing the salience of the cued referent (i.e., TOPIC CUE). To examine whether the visual cue per se modulated participants' processing, we computed a further ERP analysis time-locked to the onset of the visual cue. If yes, this should be reflected in differential, especially early sensory-evoked potentials in dependence of the CUE POSITION on screen (i.e., TOPIC LEFT, TOPIC RIGHT, and NEUTRAL BOTTOM).

6.3.4 ERP results of the visual cue per se

With the following ERP analysis we aimed to test if the implicit visual cue modulated participants' early sensory-evoked potentials (i.e., N1, P2) time-locked to the onset of the visual cue. The visual cue was presented for 66 ms and was directly followed by the pictures. However, since the timing and the position of the pictures were always the

same, early processing differences are likely to be related to the different prior CUE POSITIONS. Therefore, ERP analyses were calculated to assess the impact of CUE POSITION (i.e., TOPIC LEFT, TOPIC RIGHT, and NEUTRAL BOTTOM) in the N1 (100 to 200 ms) and P2 (250 to 350 ms) time window post onset of the visual cue and its topographical distribution by the factor ROI.[16]

As can be seen in Figure 6.3, the ERPs time-locked to the visual cue show an early negativity around 100 - 200 ms followed by a positive deflection around 250 - 350 ms. Statistical analysis in the time window of 100 - 200 ms revealed a statistically significant difference in the form of a main effect of CUE POSITION [$F(2, 56) = 9.477$, $p < .001$, $\varepsilon = 0.966$] and of an interaction of CUE POSITION X ROI [$F(16, 448) = 7.337$, $p < .001$, $\varepsilon = 0.351$]. Post hoc comparisons show that the negativity around 100 - 200 ms was strongest for the cue at the NEUTRAL BOTTOM position compared to the TOPIC LEFT and TOPIC RIGHT position. This difference was present in multiple ROIS (see Appendix D for the complete post hoc statistical results of the respective ROIS). Regarding the positivity around 250 - 350 ms, statistical analyses revealed a statistically significant main effect of CUE POSITION [$F(2, 56) = 7.628$, $p < .001$, $\varepsilon = 0.949$] and a statistically significant interaction of CUE POSITION x ROI [$F(16, 448) = 3.675$, $p = .007$, $\varepsilon = 0.255$]. Post hoc analyses show that the TOPIC LEFT cue elicited a significantly more pronounced positive deflection compared to the NEUTRAL BOTTOM cue ($p < .001$). Note, we cannot clearly disentangle the response to the cue and the response to the picture, as the pictures were always presented 66 ms after the cue. The modulation by CUE POSITION might therefore either reflect the direct effect of the cues themselves or their impact on the processing of the subsequently presented pictures.

In both cases, the results indicate that participants processed the implicit visual cues. But still, given that none of the participants noticed the presence of the cues nor was sentence processing influenced by the cue, we can assume, that the cues were processed only subconsciously.

[16] Note that for the ANOVA and post hoc *t*-tests reported in Appendix D, we report the statistical analysis based on the grand average ERPs of all available trials (i.e., 40 trials each for TOPIC LEFT vs. TOPIC RIGHT cue positions and 80 trials for the NEUTRAL BOTTOM cue position). However, conducting the analysis with a similar number of trials across conditions with random sampling of half of the NEUTRAL BOTTOM cue position trials (i.e., 40), revealed a similar pattern of results.

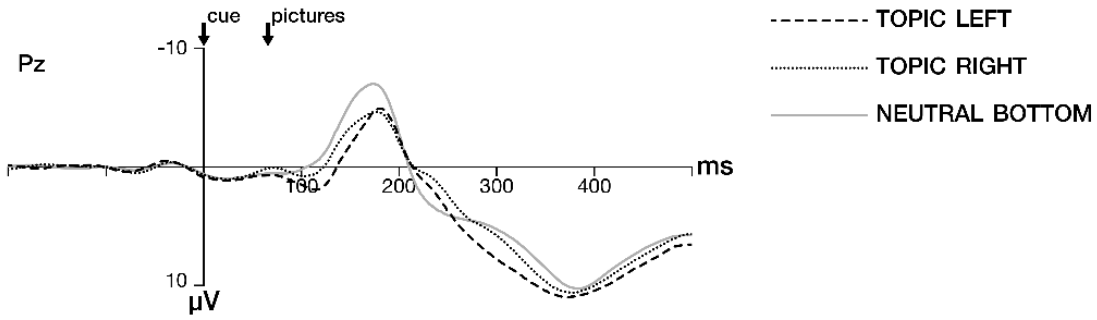

Figure 6.3: Grand average ERPs time-locked to the onset of the implicit visual cue
Grand average ERPs of one representative electrode for the factor CUE POSITION (TOPIC LEFT, TOPIC RIGHT, NEUTRAL BOTTOM) time-locked to the onset of the implicit visual cue (~66 ms cue duration directly followed by pictures). Negativity is plotted upwards.

6.3.5 Behavioural results

In the sentence-picture-verification task (in 20 % of trials) participants showed a high response accuracy across conditions indicating that participants were attentive throughout the experiment: TOPIC SO: $M = 0.91$ ($SE = 0.03$), NEUTRAL SO: $M = 0.91$ ($SE = 0.02$), TOPIC OS: $M = 0.82$ ($SE = 0.03$), NEUTRAL OS: $M = 0.81$ ($SE = 0.02$). Logit mixed models analyses of participants' response accuracy did not reveal any statistically significant differences of the fixed effects CUE ($b = 0.112$, $SE = 0.222$, $z = 0.503$, $p > .1$), WORD ORDER ($b = 0.490$, $SE = 0.339$, $z = 1.443$, $p > .1$), or the interaction of CUE x WORD ORDER ($b = 0.226$, $SE = 0.222$, $z = 1.020$, $p > .1$).

6.4 Discussion

We aimed at answering the question if an implicit visual cue to a depicted referent impacts sentence-initial processing similarly to a purely linguistic cue as revealed by ERPs (Burmester et al., 2014). With regard to the linguistic cue, the indication of the aboutness topic referent, which was subsequently mentioned in sentence-initial position, reduced the Late Positivity during online processing of OS sentences. In the present study, the linguistic topic cue was replaced by an implicit visual cue to a depicted referent in a visual scene. However, the impact of the linguistic cue on sentence-initial processing was not replicated. In fact, none of our analyses revealed

any statistically significant effects of the visual cue on sentence processing. The findings of the linguistic cue (Burmester et al., 2014) were interpreted within the Syntax-Discourse Model (Schumacher, 2014) which highlights the role of context information for meaning computation during sentence comprehension. Within this model, the impact of context information (i.e., including sentential and situational context, amongst others), is reflected in modulations of processing costs for discourse linking and updating the hitherto built mental representation of the listener. Following this model, the impact of a referent's topic-hood of the sentence-initial referent in OS sentences was attributed to reduced discourse updating costs as –in contrast to a neutral cue– the linguistic topic cue indicated one referent more salient amongst others which rendered this referent more likely to be mentioned sentence-initially. Obviously, the visual topic cue used in the present study did not increase the salience of the cued referent to a similar extent and hence did not elicit a facilitative effect on sentence processing in terms of reduced discourse updating costs.

However, the long-time neglected role of extra-linguistic (e.g., visual) information in traditional accounts of sentence comprehension (e.g., Frazier & Fodor, 1978; Friederici, 2002; Marslen-Wilson & Tyler, 1980) has been complemented by recent models underlining the close temporal integration of multimodal (e.g., visual *and* linguistic) information in the listener's current mental representation (e.g., Bower & Morrow, 1990; Zwaan, 2004). For instance, as supported by neuroscientific methods, a one-step model of sentence comprehension –integrating concomitant information from different modalities within one step– has been suggested (Hagoort & Van Berkum, 2007). More specifically, this model postulates that linguistic (for instance, sentential structure, semantics) *and* pragmatic information (from prior discourse, extra-linguistic information such as speaker's gestures or the visual world) is immediately processed by the same brain regions (namely the left inferior frontal gyrus) in order to directly map all information onto a discourse model as the basis for sentence interpretation. Similarly, the *Coordinated Interplay Account* by Crocker et al. (2010) explicitly outlines the integration of visual scene information during sentence processing, especially in cases were visual information becomes highly relevant for the interpretation and disambiguation of spoken sentences. By linking to the so called "blank screen

paradigm" in which attention shifts during sentence comprehension even occurred when depicted objects were no longer presented (Altmann, 2004), Crocker et al. (2010) suggest that mental representations of a previously presented scene still influence the comprehension process. However, the weighting of (competing) visual and linguistic information and the role of specific inherent features of visual salience cues on the strength of impact during language comprehension is not specified in any of these models.

In the following, we will discuss our findings concerning the absent visual cue impact on sentence processing against the background of previous research using different visual cues while raising possible issues with respect to the present study design. Afterwards, we will briefly discuss the word order effect, which we replicated (Section 6.4.2).

6.4.1 Null effect of visual cue on sentence-initial processing

Against the background of previous research investigating the interaction of visual cues with linguistic processing, we discuss the absent cue effect of the present study 1) with respect to different aspects of informativity of the visual scene for the comprehension process, and 2) with respect to the different impact of visual cue types (implicit vs. explicit) on sentence comprehension and production.

With respect to the informativity of the visual scene, some previous studies emphasise the need for relevance of visual information for meaning computation during language processing. For instance, listeners use depicted events –similarly to linguistic cues (i.e., case marking)– for syntactic reanalyses of locally structurally ambiguous German sentences as reflected in a reduced P600 (Knoeferle et al., 2007). Further evidence from the field of referential processing shows that visual information impacts accessibility of discourse referents only if the linguistic context is moderate, uninformative, or ambiguous (e.g., Nappa, Wessel, McEldoon, Gleitman, & Trueswell, 2009; Vogels et al., 2013). With respect to the present study design, the visual scene might not have added crucial information to the comprehension process, for instance, to assign thematic roles in the subsequent target sentence. Compared to visual scenes used in production studies (e.g., Gleitman et al., 2007; Myachykov et al., 2012), visual

scenes in the present study did *not* depict thematic role relations for the following reasons: We aimed at minimizing confounding effects of prominence-related factors known to affect participants' gaze fixations of depicted transitive events (e.g., agent-directed fixations followed by patient-directed fixations, Ganushchak, Konopka, & Chen, 2017) as well as linear ordering preferences of sentential constituents (e.g., agent precede patient theta roles, Jackendoff, 1972). In short, we tested how a single implicit cue indicating the subsequent sentence topic would affect referent accessibility and therefore we eliminated all additional factors that could have masked this single cue. Indeed, previous research shows that the more (additional) information is conveyed by the prior discourse –by the linguistic context or by depicted scenes– the more predictable are specific upcoming words as reflected in an immediate ease of sentence processing (e.g., Burmester et al., 2018; Otten & Van Berkum, 2008; Van Berkum et al., 2005). More specifically, in an experimental design rather similar to the one in the present study, we could show that depicted thematic role information of the salient referent boosted the cue-based ease of sentence processing compared to non-predictable thematic role information (Burmester et al., 2018). Moreover, for linguistic context information, Otten and van Berkum (2008) found greater priming effects of the exact discourse message than of the mere word primes. Drawing the parallel to our earlier findings, this could speak in favour of predictive processing mechanisms following the linguistic cue explicitly indicating the upcoming sentence-initial topic (Burmester et al., 2014), while with the visual cue we rather manipulated accessibility in a way similar to a word prime. Hence, maybe further depicted information such as thematic role relations would have activated further semantic features in order to constrain a more precise and coherent discourse context, and finally support predictive processing.

Moreover, with respect to the informativity of the visual scene, the simultaneous visual presence of multiple referents has been argued to reduce referent accessibility (as revealed by reduced pronoun use in, for instance, Arnold & Griffin, 2007 or Fukumura et al., 2010). In the present study, two possible referents were simultaneously depicted, which –parallel to the preceding argument– could have caused a competition of referent accessibility. This competition of accessibility was not

the case in our ERP study in which the linguistic topic cue exclusively increased the accessibility of *one* referent while not mentioning the other (i.e., *"What about the 'topic referent'?"*; Burmester et al., 2014). However, in our reading time study (Burmester et al., 2018) multiple referents were presented simultaneously and nevertheless an explicit visual (gaze) cue increased the accessibility of *one* amongst three depicted referents as reflected in sentence-initial processing ease similar to a linguistic topic cue. But, in contrast to the present study, participants in our reading time study were already familiarised with the visual scene by a multimodal lead-in *before* the gaze cue was presented. Therefore, participants might have had greater attentional capacities at the moment of processing the gaze cue than during processing the implicit cue with the subsequently presented pictures in the here presented study. Alternatively, the *type of the cue* matters, as gaze cues are more explicit and intentional in nature than the here used implicit abstract cue.

Taking previous studies using different visual cue types into account, we can assume that visual cues indeed modulate meaning computation during sentence comprehension while they seem to differ with respect to their impact on referent accessibility. As just mentioned, one type of visual cues that has been shown to clearly influence language processing, is eye gaze. This cue is a strong social-communicative cue signalling shared-attention of speaker and listener, it is hence an *intentional* cue. Crucial evidence provided by a few studies that compared both attentional and intentional cues supports the importance of the intentional component of visual cues for listeners: Similar to speaker's gaze or pointing gesture, a visual cue (i.e., black square presented for 50 ms at the location of a possible referent) influences listener's pronoun interpretation, but only if the listener was previously instructed that this abstract visual cue is intentionally created by the speaker (Nappa & Arnold, 2014). However, the same visual cue but without the previous instruction of being intentionally created by the speaker does *not* reveal an impact. Analogously to this finding, Holle et al. (2012) showed that an intentional, conversational gesture co-occurring with the speaker's speech facilitates comprehension of German ambiguous SO and OS sentences: This short hand movement (i.e., beat gesture) emphasising the subject of the sentence reduces additional processing costs for disambiguation towards the OS word order as

indicated by a reduced P600. In contrast to this gestural cue, listeners do not make use of an explicit visual attentional cue, that is, a moving point. Hence, the intentional component of visual cues plays a significant role in meaning computation during the comprehension process. This significance could be explained by assumptions of Pickering and Garrod's (2004) interactive alignment model. Accordingly, for a successful dialogue, interlocutors are assumed to develop aligned situational models. Thus, maybe intentional cues such as eye gaze offer a window into the interlocutors' mental representations in dialogue that might trigger other attentional mechanisms than those elicited by purely attentional, visual cues. However, the present study is a first step to better understand the impact of implicit attentional visual cues on sentence-initial processing, which demonstrates to us the subtle differences and difficulties of adapting study designs of sentence comprehension (Burmester et al., 2014) and sentence production (Gleitman et al., 2007).

Indeed, studying the impact of implicit visual cues on sentence comprehension might engender some caveats weakening the measurable outcome during later sentence processing. For the visual cue of the present study, we used a black square with 0.3° x 0.3° of visual angle against coloured pictures similar to the one by Gleitman et al. (2007), who used a slightly bigger cue (i.e., 0.5° x 0.5° of visual angle) against full-colour clip-arts. Moreover, Myachykov et al. (2012) used a red dot with 0.7° x 0.7° against black-white-line drawings that –similar to the cue by Gleitman et al. (2007)– influenced sentence-initial mention during production. So maybe the visual cue in our study was too subliminal and hence, its impact too short-lived to be measurable during subsequent sentence processing. Following this line of thought, modulations of participants' attention by the implicit visual cue might be less strong and less long lasting with respect to their impact on referent accessibility compared to the high accessibility degree indicated by the explicit mention of the topic in the linguistic context. This explanation is in accordance with eye-tracking data reported by Arnold and Lao (2015): Listener's trial-initial attention to depicted referents was indeed modulated by multiple visual scene-based factors, that is, different visual attentional cues, the order of the depicted referents from left to right, and listener's idiosyncratic biases. But, the attentional cues themselves did not significantly predict subsequent

pronoun interpretation and instead, the linguistic cue of sentence-initial mention was the strongest predictor. In addition, in an experimental task that requires sentence understanding, linguistic cues such as in Burmester et al. (2014) might less likely be ignored by the reader than visual cues and picture contexts such as in the present study. Hence, it is important to check whether visual cues and depicted scenes are indeed processed by the addressee of the stimuli.

However, based on the absent impact of the implicit visual cue on sentence processing in the present study, we suggest that this type of visually induced salience of discourse referents elicits a different, less strong, impact on the accessibility degree of mentally represented discourse referents compared to linguistically induced salience via topic-hood. This train of thought is supported by ERP and behavioural studies showing that linguistic stimuli (e.g., sentences, words) more strongly affect the accessibility of entities compared to pictures (e.g., Bögels, Schriefers, Vonk, & Chwilla, 2011; Fukumura et al., 2010). Moreover, a different neural processing of linguistic and visual stimuli has been suggested (e.g., Brandon & Andrew, 2007; Zhang, Begleiter, Porjesz, & Litke, 1997) which, for instance, has been explained by the more efficient semantic access and/or memory retrieval of words compared to pictures (Dorjee, Devenney, & Thierry, 2010). However, words and pictures revealed a very similar time course of the N400 (congruity) effect with differences only in the topographical distribution (Ganis, Kutas & Sereno, 1996).

In summary, reasons for the null effect of the implicit cue in our study might, on the one hand, be traced back to a weaker, less long lasting impact on referent accessibility compared to intentional or linguistic cues. On the other hand, it might to some degree be related to methodological differences to production studies in which implicit cues modulated sentence-initial mention. Disentangling these reasons and testing the impact of other types of visual attentional cues is left open for future research.

6.4.2 Word order effect

In the present study, word order effects are reflected in sustained positive deflections (at DP1 and verb position) which were more pronounced for OS compared to SO

sentences across multiple ROIs. Hence, we replicated differential processing costs for SO and OS sentences found in Burmester et al. (2014). A body of neurocognitive research concerning the processing of German sentences with varying word order demonstrates increased processing costs for OS sentences compared to their canonical (SO) counterpart as reflected in different ERP components, time windows and across different sentence positions (e.g., Holle et al., 2012; Knoeferle et al., 2007; Matzke et al., 2002; Schlesewsky et al., 2003). For this paper, the word order effect is not of primary relevance and we take it just as a "sanity check" of our ERP-data. In line with the previous literature, we argue that the processing differences between OS and SO sentences are engendered by the subject/nominative-first-preference in German leading to increased processing demands for the non-canonical and less frequent OS word order. Note that in our study the impact of word order might be confounded with the ordering of theta roles, as both grammatical role and theta role coincided in the sentence constituents, that is, subject and agent, object and patient. The increased online processing difficulties were, however, not visible in the behavioural (accuracy) results of the subsequent probe sentence-picture-verifications. In summary, we argue that the replicated word order effect during online sentence processing confirms the validity of the design: The replication of the word order effect shows that we are not dealing with a replication failure per se but rather that the different findings can be traced back specifically to cue modality.

6.5 Conclusion

All in all, previous findings speak in favour of the integration of visual and linguistic cues into listeners' mental representations –although with a different magnitude of both cue modalities. In our study, the *implicit* visual cue to a depicted referent followed by subsequent sentence-initial mention did not influence online processing of German SO and OS sentences. Hence, the impact of the linguistic topic cue on the identical sentence material could not be replicated with the present study design, although a similar type of cue was effective in previous production studies (Gleitman et al., 2007). It therefore remains an open question if comprehension and production are influenced by similar underlying processes. Hence, we conclude that the role of visual, purely

attention-directing cues for meaning computation during sentence processing needs further clarification. Future research needs to shed more light on the role of different visual cues and their interaction with intentional and attentional aspects in guiding information packaging preferences during utterance comprehension in order to disentangle experimental task specific effects.

6.6 Acknowledgements

This work was supported by the German Research Foundation (DFG) under Grant SFB 632 'Information structure'. We thank Franziska Machens and Tobias Busch for assistance in material preparation and data collection as well as Jan Ries for his help in preparation of the Figures.

6.7 Appendices

6.7.1 Appendix A

Table 6.1: Statistical results of the ERPs time-locked to DP1
Results of analysis of variance (ANOVAs) of the ERPs for the different time windows time locked to the onset of the first determiner phrase (DP1).

	df	*F*-values		
		100 - 300 ms	300 - 500 ms	500 - 700 ms
CUE (TOPIC VS. NEUTRAL)	1, 28	0.669	0.266	1.020
CUE X ROI	8, 224	0.275 (ε = 0.252)	0.575 (ε = 0.337)	0.978 (ε = 0.285)
CUE X WORD ORDER	1, 28	1.374	0.030	0.723
CUE X WORD ORDER X ROI	8, 224	1.760 (ε = 0.411)	0.564 (ε = 0.399)	1.605 (ε = 0.417)
WORD ORDER (SO VS. OS)	1, 28	0.983	1.004	6.254*
WORD ORDER X ROI	8, 224	1.872 (ε = 0.433)	1.425 (ε = 0.378)	3.004* (ε = 0.424)

Note. Greenhouse & Geisser (1959) corrected significance levels: * p <. 05; ** p <. 01; *** p <. 001. *df* = degrees of freedom. ε = Greenhouse & Geisser epsilon factor for non-spericity to adjust the original *df* according to Jennings and Wood (1976).

6.7.2 Appendix B

Figure 6.4: Grand average ERPs at selected electrodes time-locked to the onset of the target sentence.

Grand average ERPs (with baseline correction) at selected electrodes for the factors CUE (TOPIC vs. NEUTRAL) and WORD ORDER (SO vs. OS) time-locked to the onset of the target sentence, that is, the first Determiner phrase (DP1) followed by the verb and second Determiner phrase (DP2): upper panel: TOPIC SO [dotted black] vs. NEUTRAL SO [dotted grey], lower panel: TOPIC OS [solid black] vs. NEUTRAL OS [solid grey]. Negativity is plotted upwards.

6.7.3 Appendix C

Table 6.2: Statistical results of the ERPs time-locked to the implicit visual cue
Results of overall analysis of variance (ANOVAs) of the ERPs for the different time windows time-locked to the onset of the first determiner phrase (DP1) of target sentences following both cue modalities (MODALITY), that is, the implicit visual cue (i.e., VISUAL: data of the present study) and the linguistic cue (i.e., LINGUISTIC: data published in Burmester et al., 2014).

	df	*F*-values		
		100 - 300 ms	300 - 500 ms	500 - 700 ms
CUE (TOPIC VS. NEUTRAL)	1, 46	1.970	0.065	0.000
CUE x ROI	8, 368	0.311 (ε = 0.311)	0.736 (ε = 0.428)	3.454* (ε = 0.373)
CUE x WORD ORDER	1, 46	4.162*	0.440	0.829
CUE x WORD ORDER x ROI	8, 224	2.306 (ε = 0.404)	2.088 (ε = 0.452)	4.429** (ε = 0.441)
WORD ORDER (SO VS. OS)	1, 46	1.429	0.139	0.136
WORD ORDER x ROI	8, 368	2.372# (ε = 0.507)	1.666 (ε = 0.483)	1.435 (ε = 0.460)
MODALITY (VISUAL VS. LINGUISTIC) x ROI	8, 368	3.190*	1.266	1.378
MODALITY x CUE	1, 46	0.089	0.018	1.778
MODALITY x CUE x ROI	8, 368	0.538	0.592	1.891
MODALITY x WORD ORDER	1, 46	5.793*	0.567	3.717
MODALITY x WORD ORDER x ROI	8, 368	1.117	0.377	1.152
MODALITY x CUE x WORD ORDER	1, 46	0.285	0.200	3.963#
MODALITY x CUE x WORD ORDER x ROI	8, 368	0.446	1.843	1.487

Note. Greenhouse & Geisser (1959) corrected significance levels: # p <. 06; * p <. 05; ** p <. 01; *** p <. 001. *df* = degrees of freedom. ε = Greenhouse & Geisser epsilon factor for non-spericity to adjust the original *df* according to Jennings and Wood (1976).

6.7.4 Appendix D

Table 6.3: Statistical results of post hoc *t*-tests for ERPs time-locked to the implicit visual cue with respect to CUE POSITION
Results of post hoc pairwise *t*-tests of the ERPs to resolve the ROI x CUE POSITION interaction in the N1 and P2 time window (i.e., 100 - 200 ms, 250 - 350 ms) time-locked to the onset of the visual cue.

CUE POSITION	TOPIC LEFT VS. TOPIC RIGHT				TOPIC LEFT VS. NEUTRAL BOTTOM				TOPIC RIGHT VS. NEUTRAL BOTTOM			
	100 - 200 ms		250-350 ms		100-200 ms		250-350 ms		100 - 200 ms		250 - 350 ms	
ROI	*t(28)*	*p*	*t(28)*	*p*	*t(28)*	*p*	*t(28)*	*p*	*t(28)*	*p*	*t(28)*	*p*
LEFT FRONTAL	0.753	0.458	1.153	0.259	3.780	0.001*	3.726	0.001*	1.659	0.108	1.767	0.088
RIGHT FRONTAL	-1.076	0.291	-2.207	0.036	-0.344	0.734	0.909	0.371	0.914	0.368	3.131	0.004
LEFT FRONTO-CENTRAL	1.512	0.142	2.616	0.014	4.178	<.001*	3.328	0.002	2.220	0.035	1.093	0.284
RIGHT FRONTO-CENTRAL	0.237	0.814	0.289	0.775	2.573	0.016	2.907	0.007	2.247	0.033	2.755	0.010
LEFT CENTRO-PARIETAL	0.568	0.575	1.333	0.193	4.885	<.001*	2.844	0.008	4.313	<.001*	2.023	0.053
RIGHT CENTRO-PARIETAL	1.055	0.301	0.111	0.912	4.208	<.001*	1.962	0.060	3.467	0.002	2.369	0.025
FRONTAL-MIDLINE	1.542	0.134	0.804	0.428	2.523	0.018	3.068	0.005	0.397	0.694	1.954	0.061
CENTRAL MIDLINE	1.646	0.111	3.210	0.003	3.810	0.001*	4.503	<.001*	1.883	0.070	1.145	0.262
PARIETAL MIDLINE	2.254	0.032	2.625	0.014	6.538	<.001*	3.994	<.001*	5.021	<.001*	1.348	0.188

Note. Bonferroni corrected significance level (two-tailed) at α = .05: *p <. 002.

Chapter 7.

Study 3:

Sensitivity to salience: linguistic vs. visual cues affect sentence processing and pronoun resolution[17]

Language, Cognition and Neuroscience, 33(6), (2018), 784–801.

JULIANE BURMESTER[1], ANTJE SAUERMANN[2], KATHARINA SPALEK[3], AND ISABELL WARTENBURGER[1]

[1] Cognitive Sciences, Department Linguistics, University of Potsdam, Potsdam, Germany
[2] Centre for General Linguistics (ZAS), Berlin, Germany
[3] Department of German Language and Linguistics, Humboldt-Universität zu 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

[17] This chapter is adapted from: Burmester, J., Sauermann, A., Spalek, K., and Wartenburger, I. (2018). Sensitivity to salience: linguistic vs. visual cues affect sentence processing and pronoun resolution. *Language, Cognition and Neuroscience, 33*(6): 784–801. DOI: 10.1080/23273798.2018.1428758

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 into the current mental representation of the discourse.

Keywords: topic status; eye gaze; visual context; reading times; antecedent choice

7.1 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 et al., 1993; Halliday, 1968). 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; Smith & Kam, 2015; Staudte et al., 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., Bornkessel & Schlesewsky, 2006a; 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 et al., 2010) as well as production (e.g., Fukumura et al., 2010; Vogels et al., 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).[18]

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 7.1.1 and 7.1.2 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

[18] Salience has also been described in earlier frameworks, e.g., Prague`s 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).

(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, Burmester et al., 2014; Bader & Häussler, 2010; 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.

7.1.1 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[19] occur sentence-initially (e.g., Frey, 2004b; 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 et al., 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).

[19] 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).

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 & Branigan, 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, Walenski 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 et al., 2005, for German: Bosch et al., 2007; Bouma & Hopp, 2007), first-mention (e.g., Järvikivi et al., 2005), discourse prominence (e.g., topic status; Ariel, 1988; Bosch & Umbach, 2007; Colonna et al., 2014; 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 2014) 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).

7.1.2 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 et al., 2005; Knoeferle et al., 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 et al., 2007; Myachykov et al., 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).[20] 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

[20] Bock and Warren (1985) define conceptual accessibility as the ease with which a referent`s mental representation is retrieved from memory. Prat-Sala and Branigan (2000) divide 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).

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.

7.1.3 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 1.1 vs. 1.2 for the SO and OS word order).

1.1 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.´

1.2 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.´

1.3 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 1.3). 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.

7.2 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.

7.2.1 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.

7.2.2 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 Appendix A1). 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 7.1). 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 7.1 (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.

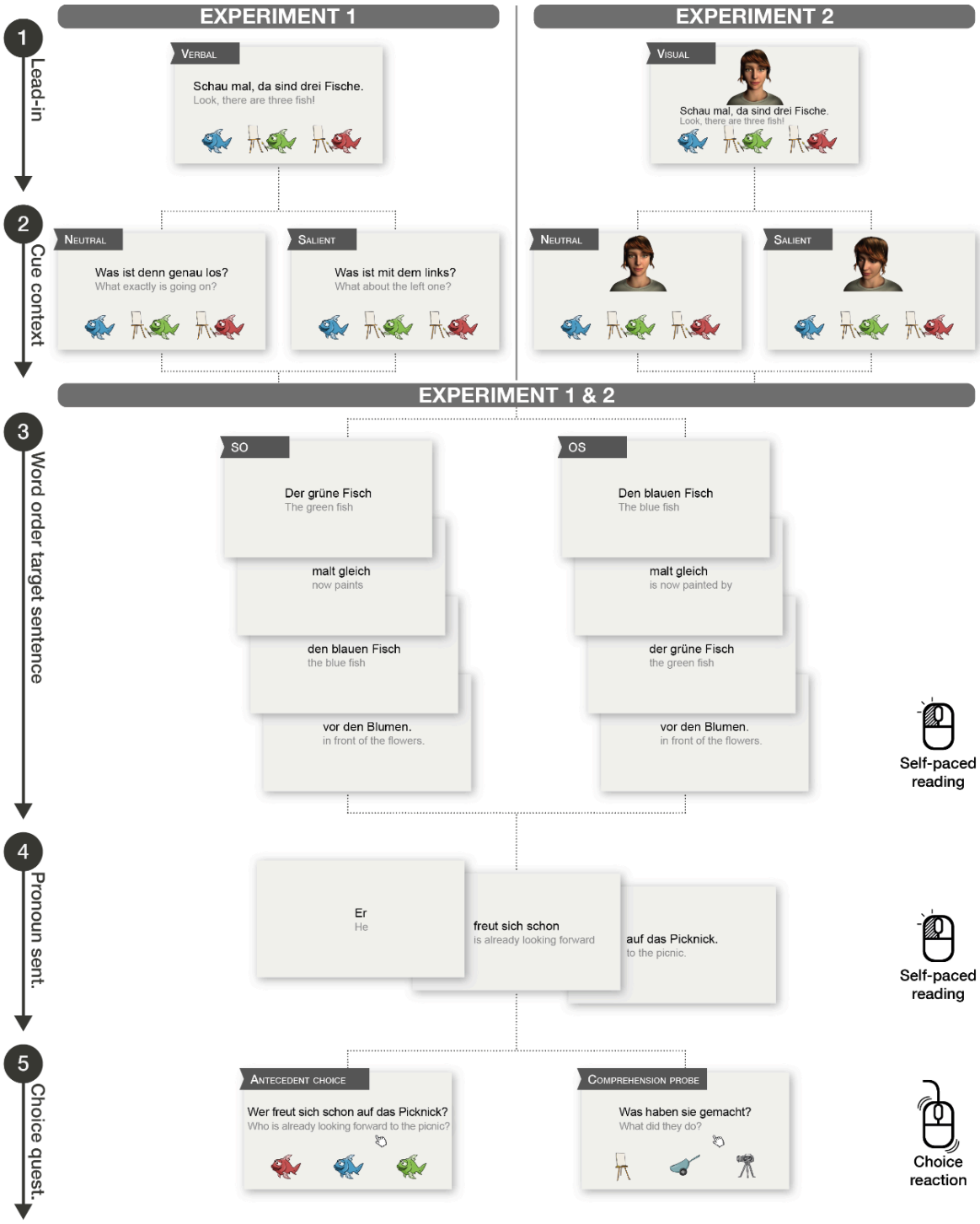

Figure 7.1: Experimental design of study 3 (explicit, intentional visual cue)
Experimental design of Exp. 1 (LINGUISTIC MODALITY) and 2 (VISUAL MODALITY) with approximate English translation written in grey. First, participants read the lead-in sentence (1). Subsequently, the salient or neutral CUE context (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.

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 7.1). 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 1.1), or object and patient (marked with accusative case (ACC), see 1.2), 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 always started with the masculine third person pronoun *"He"* (marked with nominative case

(NOM), see example 1.3) 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 7.1 (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.

7.2.3 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 7.1 (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 7.1 (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 minutes including two pauses and was followed by a short questionnaire concerning strategic behaviour.

7.2.4 Data analysis

For the statistical data analysis, linear mixed effects models were calculated using the *lme4* package, Bates et al., 2015) provided by the *R* environment (version 3.2.3., R Core Team, 2013). 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 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 7.1, 7.2, and 7.3.

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, G. E. P. & 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.

7.3 Results

Figure 7.2 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 7.3 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 7.1, 7.2, and 7.3. 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 Appendix A1.

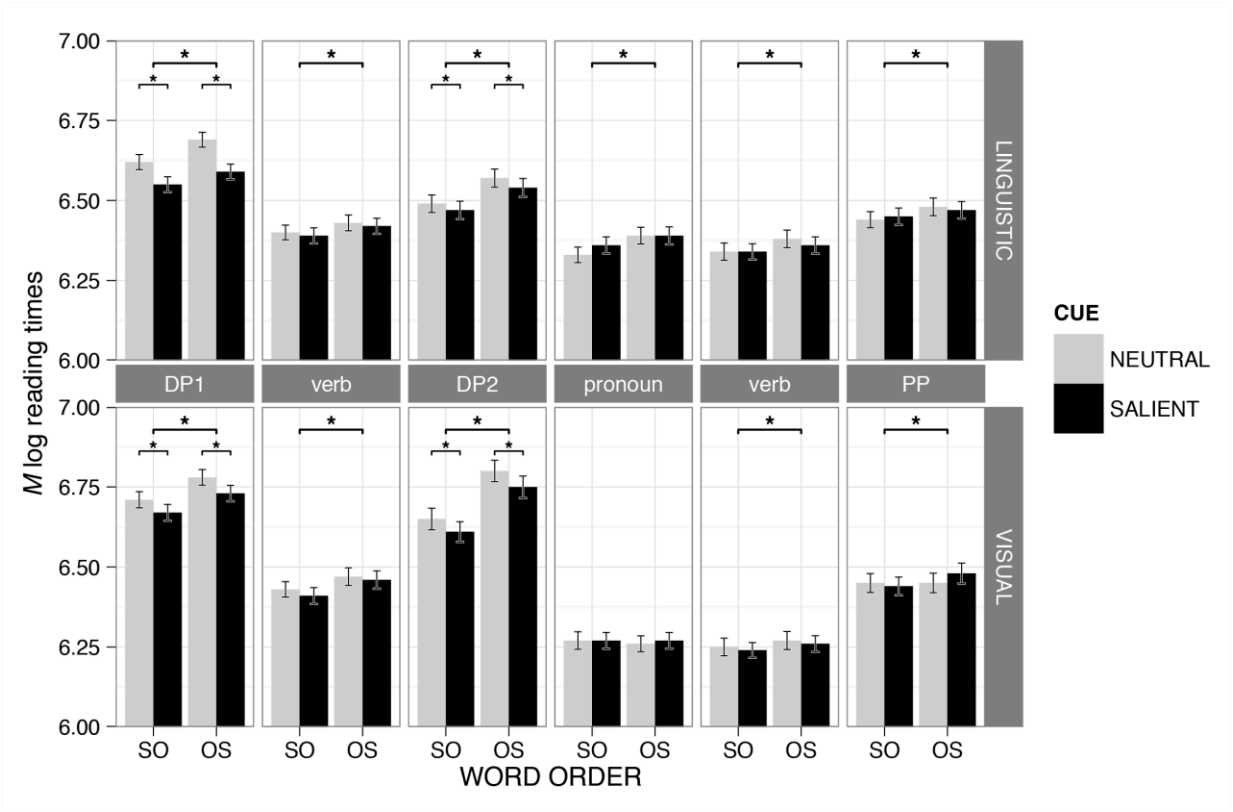

Figure 7.2: Mean *(M)* log reading times across relevant sentence positions
The Figure shows the *M* log reading times across relevant sentence positions during processing the word order sentence (DP1, verb, DP2) and the pronoun sentence (pronoun, verb, PP) for the LINGUISTIC (Exp. 1: upper panel) and 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 7.3.1).

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

7.3.1 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 7.1.

DP1. Crucially for our research question, MODALITY of cue did *not* significantly affect reading times of the 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 7.2). 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 x 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 7.2). Hence, across sentence positions WORD ORDER significantly affected sentence processing.

Table 7.1: 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 x CUE	0.042	0.026	1.60	-0.008	0.016	-0.48	-0.021	0.019	-1.10
MODALITY x WORD ORDER	0.012	0.020	0.58	0.018	0.020	0.90	0.074	0.037	**1.98***
CUE x WORD ORDER	-0.018	0.015	-1.19	-0.005	0.016	-0.34	-0.011	0.019	-0.56
MODALITY x CUE x 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.

7.3.2 Pronoun resolution

7.3.2.1 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 x WORD ORDER (see Figure 7.2 for the plotted log reading times and Table 7.2 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, 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, 2004). 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.

Table 7.2: Linear mixed effects model output for log reading times across sentence positions of the pronoun sentence

	pronoun			verb			prepositional phrase		
FIXED EFFECTS	*b*	*SE*	*z*	*b*	*SE*	*z*	*b*	*SE*	*z*
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	0.003	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 x CUE	-0.017	0.016	-1.08	0.002	0.016	0.12	0.007	0.018	0.42
MODALITY x WORD ORDER	-0.041	0.016	**-2.56***	-0.009	0.016	-0.57	-0.012	0.018	-0.69
CUE x WORD ORDER	-0.015	0.016	-0.93	-0.009	0.016	-0.53	-0.011	0.018	0.64
MODALITY x CUE x 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.

7.3.2.2 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 7.3 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 x WORD ORDER, of MODALITY x WORD ORDER, and of MODALITY x CUE x WORD ORDER (see Table 7.3 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 7.3 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.

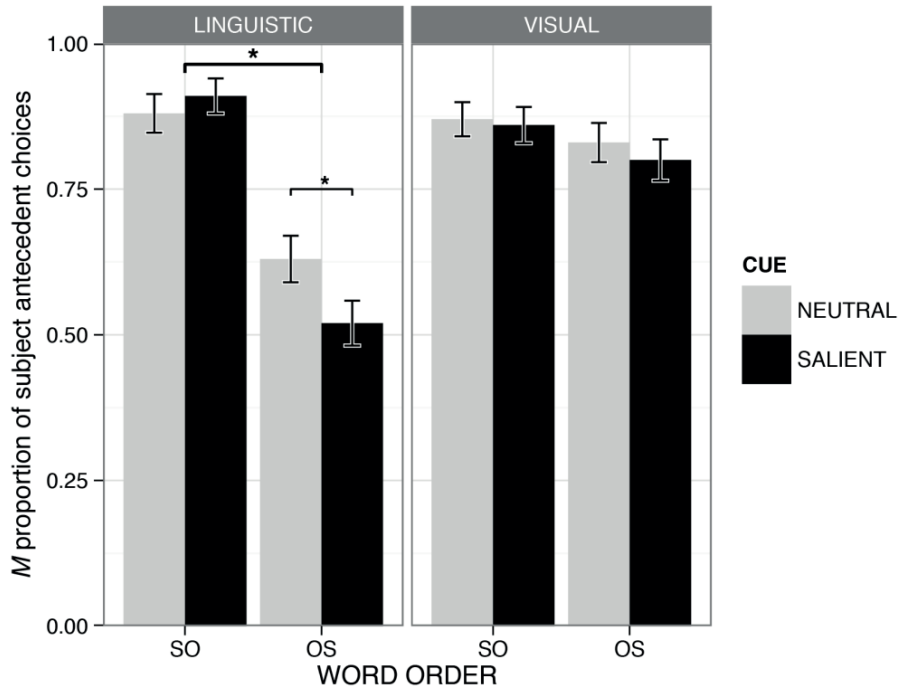

Figure 7.3: Mean *(M)* proportion of subject antecedent choices for the pronoun
M proportion of subject antecedent choices for the pronoun *"He"* for the LINGUISTIC (Exp. 1: left panel) and VISUAL MODALITY (Exp. 2: right panel). Statistically significant effects with $|z| > 1.96$ are marked with an asterisk. 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.

Table 7.3: 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 x CUE	0.078	0.212	0.37
MODALITY x WORD ORDER	1.649	0.643	**2.56***
CUE x WORD ORDER	-0.629	0.212	**-2.97****
MODALITY x CUE x 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.

7.4 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 (for details see 7.7 Supplemental results). In the following, our findings will be discussed with a focus on the impact of cue modality on sentence processing and pronoun resolution.

7.4.1 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., Burmester et al., 2014; Bornkessel & Schlesewsky, 2006a; 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 et al., 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 Appendix A1 for the results). In sum, during sentence processing we see similar facilitating effects of linguistic and visual cues.

7.4.2 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., 2014). In line with Colonna et al. (2014), 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 & 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.

7.4.3 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 referents´ 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 et al., 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 co-reference 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.

7.5 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).

7.6 Acknowledgements

This work was supported by the German Research Foundation (DFG) under Grant SFB 632 `Information structure´. AS was partially funded by the German Federal Ministry of Education and Research (BMBF) under Grant [number 01UG1411]. We thank Franziska Machens, Helene Killmer, Tobias Busch, and Jan Ries for assistance in material preparation and data collection.

7.7 Supplemental results

Results of the control factors THETA ROLE and ORIENTATION

To test if visually depicted theta role information and the position of the depicted animals in the visual scene affect sentence processing and pronoun resolution, we here report the statistical significant effects with regard to the control factors THETA ROLE (*predictable* vs. *unpredictable)* and ORIENTATION *(left-to-right vs. right-to-left).* These results were revealed by the models reported in Section 7.3. *Results* (see Table 7.1 and 7.2 for the respective final mixed effects model). For sentence processing, statistical analyses of log reading times at DP1 yielded a significant interaction of CUE x THETA ROLE (b = 0.037, SE = 0.015, t = 2.42) suggesting that the impact of the salience cue on reading times at DP1 differed with regard to the predictability of theta roles: Post hoc analyses revealed that -in comparison to the neutral cue- the impact of the salience cue was greater (i.e., DP1 was read faster) if theta roles were predictable from the scene (i.e., left or right animal cued) (b = 0.085, SE = 0.015, t = 5.61, p = .001) than if theta roles were unpredictable (i.e., central animal cued) (b = 0.048, SE = 0.015, t = 3.19, p = .010). At the following sentence position, log reading times of the verb significantly interacted with regard to CUE x WORD ORDER x ORIENTATION (b = -0.103, SE = 0.031, t = -3.27). Post hoc analyses showed that if readers were presented with animals oriented from right-to-left (i.e., left animal was patient/object) followed by a salient-first OS sentence reading times were shorter compared to a preceding neutral cue (b = 0.049, SE = 0.016, t = 3.02, p = .052). Hence, processing of OS sentences mostly benefited from the salience cue if the first-mentioned patient/object of the scene was located at the left side of the visual display. For SO sentences the ORIENTATION of the animals in the visual display did not play a significant role (p > .1). The significant interaction of CUE x THETA ROLE x ORIENTATION (b = -0.093, SE = 0.031, t = -2.96) did not reveal statistically significant results post hoc, which indicates that theta role information did not enhance processing at the position of the finite verb (p > .1). At DP2 statistical analyses revealed a significant interaction of CUE x WORD ORDER x ORIENTATION (b = -0.090, SE = 0.038, t = -2.35). Post hoc analyses showed that similar to the effect at verb position, reading times were shorter if readers were presented with animals oriented from right-to-left (i.e., left animal was patient/object) followed by a salient-first OS sentence compared to a preceding neutral cue (b = 0.062, SE = 0.020,

t = 3.15, p = .036). Again, there was no significant impact of the animals` ORIENTATION if followed by SO sentences (p > .1).

For pronoun resolution, statistical analyses of log reading times of the pronoun yielded a significant interaction of CUE x WORD ORDER x ORIENTATION (b = -0.115, SE = 0.034, t = -3.36) for which no significant differences were found in post hoc comparisons (p > .1). Regarding the spillover region, analyses of log reading times of the verb following the pronoun showed a significant interaction of CUE x THETA ROLE (b = 0.045, SE = 0.016, t = 2.76) such that the salience cue speeded reading times if theta roles were predictable from the scene (i.e., left or right animal cued) (b = 0.032, SE = 0.012, t = 2.82, p = .025), but not if theta roles were unpredictable (i.e., central animal cued) (b = -0.013, SE = 0.012, t = -1.08, p = .702). For the following prepositional phrase, the analyses revealed a significant interaction of CUE x WORD ORDER x ORIENTATION (b = -0.134, SE = 0.036, t = -3.75). Post hoc comparisons to resolve the interaction did not show a significant effect of CUE or ORIENTATION within SO or OS sentences.

Finally, both the predictability of theta roles as well as the position of pictures (i.e., left/right orientation of animals) in the visual display proved to affect processing of the word order and pronoun sentence to different dimensions. However, subject antecedent choices were not modulated by these two control factors.

References

Altmann, G.T.M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition, 93*(2), B79–B87.

American Electroencephalographic Society. (2006). Guideline 5: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology, 11,* 111–113.

American Psychological Association. (2010). *Publication Manual of the American Psychological Association*. Washington, DC, US: American Psychological Association.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics, 24*(01), 65–87.

Ariel, M. (2001). Accessibility theory: an overview. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text Representation: Linguistic and Psycholinguistic Aspects* (pp. 28–87). Amsterdam/Philadelphia: John Benjamins Publishing Company.

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

Arnold, J.E. (2010). How speakers refer: the role of accessibility. *Language and Linguistics Compass, 4*(4), 187–203.

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., Kaiser, E., Kahn, J.M., & Kim, L.K. (2013). Information structure: linguistic, cognitive, and processing approaches. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(4), 403–413.

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

Baayen, R.H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

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

Bader, M., & Häussler, J. (2010). Word order in German: a corpus study. *Lingua, 120*(3), 717–762.

Bader, M., & Meng, M. (1999). Subject-object ambiguities in German embedded clauses: an across-the-board comparison. *Journal of Psycholinguistic Research, 28*(2), 121–143.

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). Hillsdale, New York: Lawrence Erlbaum Associates.

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.

Bates, D.M., Mächler, M., & Dai, B. (2009). lme4: linear-mixed-effect models using S4 classes. Vienna, Austria: R Foundation for Statistical Computing.

Bates, D.M., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

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.

Bögels, S., Schriefers, H., Vonk, W., & Chwilla, D.J. (2011). Pitch accents in context: how listeners process accentuation in referential communication. *Neuropsychologia, 49*(7), 2022–2036.

Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A.D. (2004). Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language, 51,* 495–522.

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

Bornkessel, I., & Schlesewsky, M. (2006b). The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychological Review, 113*(4), 787–821.

Bornkessel, I., Schlesewsky, M., & Friederici, A.D. (2002). Grammar overrides frequency: evidence from the online processing of flexible word order. *Cognition, 85,* B21-B30.

Bornkessel, I., Schlesewsky, M., & Friederici, A.D. (2003). Contextual information modulates initial processes of syntactic integration: the role of inter- versus intrasentential predictions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(5), 871–882.

Bornkessel, I., Zysset, S., Friederici, A.D., Von Cramon, D.Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage, 26*(1), 221–233.

Bornkessel-Schlesewsky, I., Grewe, T., & Schlesewsky, M. (2012). Prominence vs. aboutness in sequencing: a functional distinction within the left inferior frontal gyrus. *Brain and Language, 120*(2), 96–107.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009a). 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.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009b). Minimality as vacuous distinctness: evidence from cross-linguistic sentence comprehension. *Lingua, 119*(10), 1541–1559.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2012). Linguistic sequence processing and the prefrontal cortex. *The Open Medical Imaging Journal, 6*, 47–61.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: a new dorsal–ventral stream model of sentence comprehension. *Brain and Language, 125*(1), 60–76.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology, 10,* 298.

Bornkessel-Schlesewsky, I., & Schumacher, P.B. (2016). Towards a neurobiology of information structure. In C. Féry & S. Ishihara (Eds.), *The Oxford handbook of information structure* (pp. 581–598). New York: Oxford University Press.

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.

Bradley, S.D. (2007). Dynamic, embodied, limited-capacity attention and memory: modeling cognitive processing of mediated stimuli. *Media Psychology, 9*(1), 211–239.

Brandon, A.A., & Andrew, E.B. (2007). The worth of pictures: using high density event-related potentials to understand the memorial power of pictures and the dynamics of recognition memory. *NeuroImage, 35*(1), 378–395.

Brouwer, H., Crocker, M.W., Venhuizen, N.J., & Hoeks, J.C.J. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science, 41,* 1318–1352.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research, 1446,* 127–143.

Brown, M., Savova, V., & Gibson, E. (2012). Syntax encodes information structure: evidence from on-line reading comprehension. *Journal of Memory and Language, 66*(1), 194–209.

Büring, D. (1999). Topic. In P. Bosch & R. van der Sandt (Eds.), *Studies in natural language processing. Focus: Linguistic, cognitive, and computational perspectives* (pp. 142–165). Cambridge: University Press.

Büring, D. (2011). Topic and comment. In Patrick Hogan (Ed.), *The Cambridge Encyclopedia of the Language Sciences* (p. 866). Cambridge: Cambridge University Press.

Burkhardt, P. (2005). *The Syntax-discourse interface: representing and interpreting dependency*. John Benjamins Publishing Company.

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: evidence from event-related brain potentials. *Brain and Language, 98*(2), 159–168.

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *Cognitive Neuroscience and Neuropsychology, 18*(17), 1851–1854.

Burkhardt, P., & Roehm, D. (2007). Differential effects of saliency: an event-related brain potential study. *Neuroscience Letters, 413*(2), 115–120.

Burmester, J., Sauermann, A., Spalek, K., & Wartenburger, I. (2018). Sensitivity to salience: linguistic vs. visual cues affect sentence processing and pronoun resolution. *Language, Cognition and Neuroscience, 33*(6), 784–801.

Burmester, J., Spalek, K., & Wartenburger, I. (2014). Context updating during sentence comprehension: the effect of aboutness topic. *Brain and Language, 137,* 62–76.

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* (PhD thesis). Universität Potsdam, Potsdam.

Choi, H.-W. (1996). *Optimizing structure in context: scrambling and information structure* (PhD thesis). Stanford University.

Clark, H., & Haviland, S.E. (1977). Comprehension and the given-new contract. In R.O. Freedle (Ed.), *Discourse production and comprehension* (pp. 1–40). Hillsday, New Jersey: Erlbaum.

Clark, H.H., & Clark, E.V. (1977). *Psychology and language: an introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

Cohn, N., Paczynski, M., Jackendoff, R., Holcomb, P.J., & Kuperberg, G.R. (2012). (Pea)nuts and bolts of visual narrative: structure and meaning in sequential image comprehension. *Cognitive Psychology, 65*(1), 1–38.

Colonna, S., Schimke, S., & Hemforth, B. (2014). Information structure and pronoun resolution in German and French: evidence from the visual-world paradigm. In B. Hemforth, B. Mertins, & C. Fabricius-Hansen (Eds.), *Studies in theoretical psycholinguistics: Vol. 44. Psycholinguistic approaches to meaning and understanding across languages* (pp. 175–195). Dordrecht: Springer.

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.

Corbetta, M., & Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Coulson, S. (2006). Electrifying results ERP data and cognitive linguistics. In M. Gonzales-Marquez, I. Mittelberg, S. Coulson, & M. Spivey (Eds.), *Methods in Cognitive Linguistics* (pp. 401–423). Amsterdam: John Benjamins Publishing Company.

Coulson, S., Federmeier, K.D., van Petten, C., & Kutas, M. (2005). Right hemisphere sensitivity to word- and sentence-level context: evidence from event-related rain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 129–147.

Coulson, S., King, J.W., & Kutas, M. (1998). Expect the unexpected: event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*(1), 21–58.

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

Cowan, N., Elliott, E.M., Scott Saults, J., Morey, C.C., Mattox, S., Hismjatullina, A., & Conway, A.R.A. (2005). On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*(1), 42–100.

Cowles, H.W. (2003). *Processing information structure: evidence from comprehension and production* (PhD thesis). University of California, San Diego.

Cowles, H.W. (2007). The influence of "aboutness" on prominal coreference. *ZAS Papers in Linguistics,* 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.

Cowles, H.W., Kluender, R., Kutas, M., & Polinsky, M. (2007). Violations of information structure: an electrophysiological study of answers to wh-questions. *Brain and Language, 102*(3), 228–242.

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.

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

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.P.G. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eyemovements: a window on mind and brain* (pp. 471–486). Elsevier: Oxford.

Delogu, F., Drenhaus, H., & Crocker, M.W. (2018). On the predictability of event boundaries in discourse: an ERP investigation. *Memory & Cognition, 46*(2), 315–325.

Diedrichsen, E. (2008). Where is the precore slot? Mapping the layered structure of the clause and German sentence topology. In R.D. Van Valin, Jr. (Ed.), *Investigations of the Syntax–Semantics–Pragmatics Interface* (pp. 203–224). Amsterdam: John Benjamins Publishing Company.

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

Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*(03), 357.

Dorjee, D., Devenney, L., & Thierry, G. (2010). Written words supersede pictures in priming semantic access: a P300 study. *NeuroReport, 21*(13), 887–891.

Dryer, M.S. (2013). Order of subject, object and verb. In M.S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online.* (Available online at http://wals.info/chapter/81). Leipzig: Max Planck Institute for Evolutionary Anthropology.

Dunn, B.R., Dunn, D.A., Languis, M., & Andrews, D. (1998). The relation of ERP components to complex memory processing. *Brain and Cognition, 36,* 355–376.

Evans, K.M., & Federmeier, K.D. (2007). The memory that's right and the memory that's left: event-related potentials reveal hemispheric asymmetries in the encoding and retention of verbal information. *Neuropsychologia, 45*(8), 1777–1790.

Fanselow, G., & Lenertová, D. (2011). Left peripheral focus: mismatches between syntax and information structure. *Natural Language & Linguistic Theory, 29*(1), 169–209.

Federmeier, K.D., & Kutas, M. (2001). Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: General, 27*(1), 202–224.

Féry, C., & Krifka, M. (2008). Information structure: notional distinctions, ways of expression. In P. van Sterkenburg (Ed.), *Unity and diversity of languages* (pp. 123–136). Amsterdam: John Benjamins Publishing Company.

Fiebach, C.J., Schlesewsky, M., & Friederici, A.D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German wh-questions. *Journal of Memory and Language, 47,* 250–272.

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, Italy: Martello-Giunti.

Frazier, L., & Fodor, J.D. (1978). The sausage machine: a new two-stage parsing model. *Cognition, 6*(4), 291–325.

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

Frey, W. (2004a). The grammar-pragmatics interface and the German prefield. *Sprache Und Pragmatik, 52,* 1–39.

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

Frey, W. (2006). Contrast and movement to the German prefield In V. Molnár (Ed.), *The Architecture of Focus: Studies in Generative Grammar (SGG)* (pp. 235–264). Berlin: Walther de Gruyter.

Frey, W. (2007). Some contextual effects of aboutness topics in German. In A. Späth (Ed.), *Interfaces and interface conditions: language, context and cognition* (pp. 329–348). Berlin: Walther de Gruyter.

Friederici, A.D., Wang, Y., Herrmann, C.S., Maess, B., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: a magnetoencephalographic study. *Human Brain Mapping, 11*, 1–11.

Friederici, A.D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84.

Friederici, A.D., Steinhauer, K., Mecklinger, A., & Meyer, M. (1998). Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. *Biological Psychology, 47*(3), 193–221.

Friedrich, C.K., & Kotz, S.A. (2007). Event-related potential evidence of form and meaning coding during online speech recognition. *Journal of Cognitive Neuroscience, 19*(4), 594–604.

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition, 85,* B83-B92.

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

Ganis, G., Kutas, M., & Sereno, M.I. (1996). The search for "common sense": an electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience, 8*(2), 89–106.

Ganushchak, L.Y., Konopka, A.E., & Chen, Y. (2017). Accessibility of referent information influences sentence planning: an eye-tracking study. *Frontiers in Psychology, 8,* 250.

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.

Gernsbacher, M.A., & Hargreaves, D.J. (1988). Accessing sentence participants: the advantage of first mention. *Journal of Memory and Language, 27,* 699–717.

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

Givón, T. (1988). The pragmatics of word-order: predictability, importance and attention. In M. Hammond, E.A. Moravczik, & J.R. Wirth (Eds.), *Studies in syntactic typology* (pp. 243–284). 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.

Gorrell, P. (2000). The subject-before-object-preference in German clauses. In B. Hemforth (Ed.), *German sentence processing* (pp. 25–63). Dordrecht: Kluwer Academic Publishers.

Gratton, G., Coles, M.G.H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55,* 468–484.

Greenhouse, S.W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*(2), 95–112.

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., Cramon, D.Y. von, & Schlesewsky, M. (2006). Linguistic prominence and Broca's area: the influence of animacy as a linearization principle. *NeuroImage, 32*(3), 1395–1402.

Grosz, B.J., & Sidner, C.L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics, 12*(3), 175–204.

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

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

Gundel, J.K. (1988). Universals of topic-comment structure. In M. Hammond, E.A. Moravczik, & J.R. Wirth (Eds.), *Studies in syntactic typology* (pp. 209–239). Amsterdam: John Benjamins Publishing Company.

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

Hagoort, P. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*(4), 439–483.

Hagoort, P., & Van Berkum, J.J.A. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1481), 801–811.

Haider, H., & Rosengren, I. (1998). Scrambling. In I. Rosengren (Ed.), *Netzwerk Sprache und Pragmatik* (pp. 1–104). Lund: Germanistisches Institut der Univeristät Lund.

Halliday, M.A.K. (1968). Notes on transitivity and theme in English. *Journal of Linguistics, 4*(2), 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.

Hasson, U., & Giora, R. (2006). Experimental methods for studying the mental representation of language. In M. Gonzales-Marquez, I. Mittelberg, S. Coulson, & M. Spivey (Eds.), *Methods in Cognitive Linguistics* (pp. 302–322). Amsterdam: John Benjamins Publishing Company.

Haviland, S.E., & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13*(5), 512–521.

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.

Hemforth, B. (1993). *Kognitives Parsing: Repräsentation und Verarbeitung sprachlichen Wissens* (PhD thesis). University of Bochum, Bochum.

Hirotani, M., & Schumacher, P.B. (2011). Context and topic marking affect distinct processes during discourse comprehension in Japanese. *Journal of Neurolinguistics, 24*(3), 276–292.

Hockett, C.F. (1958). *A course in modern linguistics*. New York: McMillan.

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.

Hung, Y.-C., & Schumacher, P.B. (2012). Topicality matters: position-specific demands on Chinese discourse processing. *Neuroscience Letters, 511*(2), 59–64.

Hung, Y.-C., & Schumacher, P.B. (2014). Animacy matters: ERP evidence for the multi-dimensionality of topic-worthiness in Chinese. *Brain Research, 1555,* 36–47.

Jackendoff, R. (1972). *Semantic interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.

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

Jacobs, J. (2001). The dimensions of topic–comment. *Linguistics, 39*(4), 641–681.

Jacobs, J. (1988). Fokus-Hintergrund-Gliederung und Grammatik. In H. Altmann (Ed.), *Intonationsforschungen* (pp. 89–134). Tübingen: Niemeyer.

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.

Jefferies, E., Lambon Ralph, M.A., & Baddeley, A.D. (2004). Automatic and controlled processing in sentence recall: the role of long-term and working memory. *Journal of Memory and Language, 51*(4), 623–643.

Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). New York: Routledge.

Jennings, R., & Wood, Ch. C. (1976). The *ε*-adjustment procedure for repeated-measures analyses of variance. *Psychophysiology, 13*(3): 277–278.

Johnson-Laird, P.N. (1980). Mental models in cognitive science. *Cognitive Science Proceedings, 4,* 71–115.

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

Kaan, E., Dallas, A.C., & Barkley, C.M. (2007). Processing bare quantifiers in discourse. *Brain Research, 1146,* 199–209.

Kaan, E., Harris, A., Gibson, E., & Holocomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*(2), 159–201.

Kaiser, E. (2006). Effects of topic and focus on salience. In Christian Ebert, Cornelia Endriss (Ed.), *Proceedings of Sinn und Bedeutung 10* .

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

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.

Kamide, Y., Altmann, G.T.M., & 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.

Kempen, G., & Harbusch, K. (2005). The relationship between grammaticality ratings and corpus frequencies: a case study into word order variability in the midfield of German clauses. In S. Kepser & M. Reis (Eds.), *Linguistic evidence – empirical, theoretical, and computational perspectives* . Berlin: Mouton De Gruyter.

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.

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.

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

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica, 55*(3), 243–276.

Kristensen, L.B., Engberg-Pedersen, E., Højlund Nielsen, A., & Wallentin, M. (2013). The influence of context on word order processing – an fMRI study. *Journal of Neurolinguistics, 26*(1), 73–88.

Kristensen, L.B., Engberg-Pedersen, E., & Poulsen, M. (2014). Context improves comprehension of fronted objects. *Journal of Psycholinguistic Research, 43*(2), 125–140.

Kristensen, L.B., Engberg-Pedersen, E., & Wallentin, M. (2014). Context predicts word order processing in Broca's region. *Journal of Cognitive Neuroscience, 26*(12), 2762–2777.

Kutas, M., & Federmeier, K.D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*(1), 621–647.

Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205.

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

Lau, E.F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience, 9*(12), 920–933.

Lenerz, J. (1977). *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Narr.

Lenerz, J. (2000). Word order variation: competition or co-operation? In G. Müller & W. Sternefeld (Eds.), *Competition in Syntax.* (pp. 249–282). Berlin & New York: Walther de Gruyter.

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

Luck, S.J., & Hillyard, S.A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31,* 291–308.

MacWhinney, B. (1977). Starting points. *Language, 53,* 152–168.

Mangun, G.R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32,* 4–18.

Marslen-Wilson, W., & Tyler, L.K. (1980). The temporal structure of spoken language understanding. *Cognition, 8*(1), 1–71.

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.

McKoon, G., Ward, G., & Ratcliff, R. (1993). Morphosyntactic and pragmatic factors affecting the accessibility of discourse entities. *Journal of Memory and Language,* 56–75.

Meng, M., Bader, M., & Bayer, J. (1999). Sprachverstehen im Kontext: Zur Rolle struktureller und semantisch-pragmatischer Information bei der Auflösung von Subjekt-Objekt-Ambiguitäten im Deutschen. In I. Wachsmuth & B. Jung (Eds.), *Proceedings der 4. Fachtagung der Gesellschaft für Kognitionswissenschaft* (pp. 244–249). St. Augustin, Germany: Infix.

Mitchell, D.C. (2004). On-line methods in language processing: introduction and historical review. In M. Carreiras & C. Clifton, JR. (Eds.), *The on-line study of sentence comprehension: eyetracking, ERPs and beyond* (pp. 15–32). New York: Psychology Press.

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

Müller, G. (1999). Optimality, markedness, and word order in German. *Linguistics, 37,* 777–818.

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.

Nagy, M.E., & Rugg, M.D. (1989). Modulation of event-related potentials by word repetition: The effects of inter-item lag. *Psychophysiology, 26*(4), 431–436.

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.

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.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J.D. (2005). Decision making, the P3, and the locus coeruleus--norepinephrine system. *Psychological Bulletin, 131*(4), 510–532.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh handedness inventory. *Neuropsychologia, 9,* 97–113.

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

Osterhout, L., & Holcomb, P.J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language,* 785–806.

Otten, M., & Van Berkum, J.J.A. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Research, 1153,* 166–177.

Otten, M., & Van Berkum, J.J.A. (2008). Discourse-based word anticipation during language processing: prediction or priming? *Discourse Processes, 45*(6), 464–496.

Penney, C.G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17*(4), 398–422.

Pickering, M.J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Science, 27,* 169–226.

Pittner, K., & Berman, J. (2008). *Deutsche Syntax* (3rd ed.). Tübingen: Gunter Narr Verlag.

Portner, P. (2007). Instructions for interpretation as separate performatives. In Kerstin Schwabe & Susanne Winkler (Eds.), *On Information Structure, Meaning and Form: Generalizations across languages.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

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.

Prince, E.F. (1981). Topicalization, focus-movement, and yiddish-movement: a pragmatic differentiation. In Berkley Linguistics Society (Ed.), *Proceedings of the seventh annual meeting of the Berkeley Linguistics Society* (pp. 249–264).

Prince, E.F. (1992). The ZPG letter: subjects, definiteness, and information-status. In S.A. Thompson & W.C. Mann (Eds.), *Discourse description: Diverse analyses of a fund raising text* (pp. 295–325). Amsterdam: John Benjamins Publishing Company.

R Core Team. (2013). *R: a language and environment for statistical computing*, Vienna, Austria.

Reinhart, T. (1981). Pragmatics and linguistics: an analysis of sentence topics. *Philosophica, 27*(1), 53–94.

Roehm, D., Bornkessel-Schlesewsky, I., Rösler, F., & Schlesewsky, M. (2007). To predict or not to predict: influences of task and strategy on the processing of semantic relations. *Journal of Cognitive Neuroscience, 19*(8), 1259–1274.

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

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

Rösler, F., Pechmann, T., Streb, J., Röder, B., & Henninghausen, E. (1998). Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory and Language, 38,* 150–176.

Rossi, S., Jürgenson, I.B., Hanulíková, A., Telkemeyer, S., Wartenburger, I., & Obrig, H. (2011). Implicit processing of phonotactic cues: evidence from electrophysiological and responses. *Journal of Cognitive Neurosciene, 23*(7), 1752–1764.

Rugg, M.D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology, 22*(6).

Scheepers, C., & Crocker, M.W. (2004). Constituent order priming from reading to listening: a visual-world study. In M. Carreiras & C. Clifton, JR. (Eds.), *The on-line study of sentence comprehension: eyetracking, ERPs and beyond* (pp. 167–185). New York: Psychology Press.

Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in German. In B. Hemforth (Ed.), *German sentence processing* (pp. 65–93). Dordrecht: Kluwer Academic Publishers.

Schlesewsky, M., Bornkessel, I., & Frisch, S. (2003). The neurophysiological basis of word order variations in German. *Brain and Language, 86*(1), 116–128.

Schumacher, P.B. (2014). Content and context in incremental processing: "the ham sandwich" revisited. *Philosophical Studies, 168*(1), 151–165.

Schumacher, P.B., & Hung, Y.-C. (2012). Positional influences on information packaging: insights from topological fields in German. *Journal of Memory and Language, 67*(2), 295–310.

Schütze, C., & Sprouse, J. (2013). Judgment data. In R.J. Podesva & D. Sharma (Eds.), *Research methods in linguistics* (pp. 27–50). Cambridge: Cambridge University Press.

Schwarzschild, R. (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics, 7,* 141–177.

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

Siewierska, A. (1993). On the interplay of factors in the determination of word order. In J. Jacobs, A. von Stechow, W. Sternefeld, & T. Vennemann (Eds.), *Syntax: an international handbook of contemporary research* (pp. 826–846). Berlin: Walther de Gruyter.

Skopeteas, S., & Fanselow, G. (2010). Effects of givenness and constraints on free word order. In M. Zimmermann & C. Féry (Eds.), *Oxford linguistics. Information structure: theoretical, typological, and experimental perspectives* (pp. 307–331). Oxford, New York: Oxford University Press.

Skopeteas, S., & Féry, C. (2007). Contrastive topic in pairing answers: a cross-linguistic production study. In S. Featherston & W. Sternefeld (Eds.), *Linguistic Evidence 2006: Linguistics in Search of its Evidential Base* [Studies in Generative Grammar] (pp. 327–347). Berlin: Mouton De Gruyter.

Skopeteas, S., Fiedler, I., Hellmuth, S., Schwarz, A., Stoel, R., Fanselow, G., Krifka, M. (2006). Questionnaire on information structure (QUIS): Reference manual (ISIS | Working Papers of the SFB 632). Universität Potsdam: Universitätsverlag Potsdam.

Smith, W.G., & Kam, C.L.H. (2015). Children's use of gesture in ambiguous pronoun interpretation. *Journal of Child Language, 42*(03), 591–617.

Spevack, S.C., Falandays, J.B., Batzloff, B., & Spivey, M.J. (2018). Interactivity of language. *Language and Linguistics Compass, 12*(7), e12282.

Speyer, A. (2004). Competing constraints on Vorfeldbesetzung in German. In B. Shaer, W. Frey, & C. Maienborn (Eds.), *Proceedings of the dislocated elements Workshop. ZASPiL 35* (pp. 519–541).

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 Company.

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

Stalnaker, R. (1974). Pragmatic presuppositions. In Milton K. Munitz & Peter K. Unger (Ed.), *Semantics and philosophy* (pp. 197–214). New York: New York University Press.

Staudte, M., & Crocker, M.W. (2011). Investigating joint attention mechanisms through spoken human–robot interaction. *Cognition, 120*(2), 268–291.

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.

Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic Inquiry, 31*(4), 649–689.

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.

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

Stolterfoht, B., Friederici, A., Alter, K., & Steube, A. (2007). Processing focus structure and implicit prosody during reading: differential ERP effects. *Cognition, 104*(3), 565–590.

Sutton, S., Braren, M., Zubin, J., & John, E.R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150*(3700), 1187–1188.

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.

Tomlin, R.S. (1986). *Basic word order. Functional principles*. London: Croom Helm.

Tomlin, R.S. (1995). Focal attention, voice, and word order: an experimental, cross-linguistic study. In P. Downing & M. Noonan (Eds.), *Word order in discourse* (pp. 517–555). Amsterdam and Philadelphia: John Benjamins Publishing Company.

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

Turner, E.A., & Rommetveit, R. (1968). Focus of attention in recall of active and passive sentences. *Journal of Verbal Learning and Verbal Behavior, 7*(2), 543–548.

Vallduvi, E., & Engdahl, E. (1996). The linguistic realisation of information packaging. *Linguistics, 34,* 459–519.

Van Berkum, J.J.A. (2004). Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things? In M. Carreiras & C. Clifton, JR. (Eds.), *The on-line study of sentence comprehension: eyetracking, ERPs and beyond* (pp. 229–270). New York: Psychology Press.

Van Berkum, J.J.A. (2008). Understanding sentences in context: What brain waves can tell us. *Current Directions in Psychological Science, 17*(6), 376–380.

Van Berkum, J.J.A. (2009). The neuropragmatics of ´simple´ utterance comprehension: an ERP review. *Semantic and Pragmatics: from Experiment to Theory,* 276–316.

Van Berkum, J.J.A. (2012). The electrophysiology of discourse and conversation. In M. Spivey, M. Joanisse, & K. McRae (Eds.), *The Cambridge handbook of psycholinguistics* (pp. 589–613). Cambridge: Cambridge University Press.

Van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 443–467.

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.

Van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Science, 3*(2), 131–150.

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

Verleger, R. (1988). Event-related potentials and cognition: A critique of the context updating hypothesis and an alternative interpretation of P3. *Behavioral and Brain Sciences, 11*(3), 343.

Vogels, J. (2014). *Referential choices in language production: The role of accessibility* (PhD thesis). Tilburg University, Tilburg.

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.

Vogels, J., Krahmer, E., & Maes, A. (2014). How cognitive load influences speakers' choice of referring expressions. *Cognitive Science, 39*(6), 1396-1418.

Wang, L., Schlesewsky, M., Philipp, M., & Bornkessel-Schlesewsky, I. (2012). The role of animacy in online argument interpretation in Mandarin Chinese. In M. Lamers & P.d. Swart (Eds.), *Case, word order and prominence* (pp. 91–119). Dordrecht: Springer Netherlands.

Wang, L., & Schumacher, P.B. (2013). New is not always costly: evidence from online processing of topic and contrast in Japanese. *Frontiers in Psychology, 4*(363).

Weber, A., & Müller, K. (2004). Word order variation in German main clauses: a corpus analysis. In *Proceedings of the 20th international conference on computational linguistics* (pp. 71–77). Geneva, Switzerland.

Weskott, T., Stolterfoht, B., Bornkessel, I., & Schlesewsky, M. (2004). *The task-dependency of acceptability judgements: processing scrambling and topicalization in German*. Paper presented at the 26th DGfS, Johannes-Gutenberg-Universität, Mainz.

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

Willems, R.M., Özyürek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Science, 20*(7), 1235–1249.

Wolff, S., Schlesewsky, M., Hirotani, M., & Bornkessel-Schlesewsky, I. (2008). The neural mechanisms of word order processing revisited: electrophysiological evidence from Japanese. *Brain & Language, 107,* 133–157.

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. 593–2598). Japan: Cognitive Science Society.

Zhang, X.L., Begleiter, H., Porjesz, B., & Litke, A. (1997). Visual object priming differs from visual word priming: an ERP study. *Electroencephalography and Clinical Neurophysiology, 102*(3), 200–215.

Zwaan, R.A. (2004). The immersed experiencer: towards an embodied theory of language comprehension. In B.H. Ross (Ed.), *The psychology of learning and motivation: advances in research and theory* (pp. 35–62). New York: Acadamic Press.

Zwaan, R.A., & Radvansky, G.A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162–185.

Erklärung

Hiermit erkläre ich, Juliane Burmester, dass ich die vorliegende Dissertation selbstständig verfasst habe und über die Beiträge meiner Koautorinnen hinaus (wie in der beiliegenden Erklärung über die Beiträge zur Gemeinschaftsveröffentlichungen beschrieben) keine Hilfe Dritter in Anspruch genommen habe. Bei der Abfassung der Arbeit wurden alle Regelungen guter wissenschaftlicher Standards eingehalten. Ich erkläre ferner, dass ich diese Arbeit ohne Benutzung anderer als der angegeben Quellen und Hilfsmittel verfasst habe.

Götz, den 1. Juli 2019

Juliane Burmester