Universität zu Köln
Philosophische Fakultät
Institut für Linguistik und Phonetik

Masterarbeit zur Erlangung des akademischen Grades

MASTER OF ARTS

im Fach Linguistik mit dem Profil Phonetik

The influence of prosodic, syntactic and lexical aspects on referent tracking in German

Erstgutachter: Apl. Prof. Dr. Stefan Baumann

Zweitgutachter: Prof. Dr. Klaus von Heusinger

Vorgelegt von

Hae-Eun Cho

Abgegeben am

07.04.2020

Abstract

This thesis investigates the influence of syntactic, semantic, prosodic and information structural factors on the referential bias in a highly controlled story continuation task in spoken German.

Keywords: Prosodic prominence, discourse prominence, givenness, focus, implicit causality bias, story continuation task.

Contents

List of Figures

List of Tables

1 Introduction

Everyday conversation involves a constant exchange of information between a speaker and a listener. Information is often conveyed in terms of entities, i.e. people or objects, which are denoted by linguistic expressions. The term *referent* is used to describe the cognitive representation of an entity which is denoted by the linguistic expression (Lambrecht 1994). *Reference* describes the relation between linguistic expressions, on the one hand, and their cognitive representations, on the other hand.

Throughout a *discourse*, speakers often use varying linguistic expressions in referring back to the same referent. This is because the speaker takes into consideration the cognitive state of the listener when planning his or her speech. The speaker delivers the information in such a way that the listener requires the least amount of cognitive effort in processing the information. In an ideal world, the listener is able to correctly identify all referents the speaker thinks of at the specific moment in discourse. In the real world, however, speakers and listeners are not synced with each other perfectly. In contrast to the speaker, who knows who is being referred to by the pronoun *he* in (1), the listener might not be able to reliably identify the correct referent to the pronoun based on the information available in (1) alone.

(1) John$_{i}$ saw Harry$_{j}$. He$_{i/j}$ was on the way to the store.

This thesis investigated the influence of different linguistic factors on the listener's choice of reference in spoken discourse. Different linguistic factors, including levels of syntax, semantics, prosody and information structure, were incorporated in a highly controlled story continuation task in German. The aim of this study was to establish those linguistic factors which are able to predict the preference for one referent over another. Another objective of this thesis was to examine whether some of the linguistic factors outweighed others and whether interaction between the different linguistic factors would lead to the amplification or nullification of certain factors. The choice of the linguistic factors were motivated by Kaiser (2011) and Kehler et al. (2008). Kaiser (2011) investigated the interpretation of pronouns in the subsequent discourse by conducting visual eye-tracking experiments. Although Kaiser (2011) used auditory stimuli in her experiments, their prosodic realisation were not controlled for sufficiently. The stimuli were merely described to be realised either with "normal intonation, without any special accenting on the pronoun" or simply with "contrastive pitch" (Kaiser 2011: 1649). Additional information on the prosodic realisation was not provided. For this study, the

prosodic realisation of the items was controlled for in terms of different accent types which signalled different information structural interpretations. Different accent types were chosen in order to evoke processing differences on the part of the listener.

(2) a. John$_{i}$ annoyed Mary$_{j}$ because <u>he</u> sang loudly.
b. John$_{i}$ admired Mary$_{j}$ because <u>she</u> sang beautifully.

Kehler et al. (2008) examined the preference for two competing referents by conducting story continuation tasks.[1] Their findings rejected the existence of a heuristic subject bias, which was proposed by Grosz et al. (1995), in favour of an implicit causality bias. Implicit causality is a semantic property of transitive verbs which attributes different types of explanations systematically to one of its binding arguments (Bott & Solstad 2014). This is exemplified in (2). While *annoyed* tends to assign the explanation to the first referent, *John*, *admired* tends to assign the explanation to the second referent, *Mary*.

In the following, §2 will clarify which of the many information structural definitions of *givenness*, *topic* and *focus* will be used in this thesis. Then, the basic literature on reference tracking in discourse, including Centering Theory (Grosz et al. 1995) and different prominence hierarchies (Ariel 1991; Givón 1983; Gundel et al. 1993), will be covered. A selected number of studies on prosodic prominence in German will be discussed as well. §2 will be concluded by introducing a new and promising account on implicit causality. §3 will detail the story continuation task conducted for this thesis, including the research question and hypotheses, the method and statistical analyses as well as the results. §4 will discuss the results of the story continuation task with regard to the theoretical background introduced in §2. §5 will conclude this thesis with a summary and an outlook on future research.

2 Reference tracking in discourse

2.1 Information structure

Everyday conversation involves a constant exchange of information between two or more interlocutors. The exchange of information between the interlocutors, usually a speaker and a listener, can either be concerned with the content of the information itself or the way the information is conveyed.

[1] Kehler et al. (2008) use the term *passage-completion experiment*.

Information structure (Halliday 1967), also known as *information packaging* (Chafe 1976), is concerned with the latter, i.e. the way a speaker packages the information he or she wants to deliver to the listener.

Information that is known to be shared by the speaker and the listener is called *common ground* (Stalnaker 1974; Karttunen 1974; Lewis 1979). In the course of a conversation, the common ground constantly changes as new information is introduced to it. Thus the speaker is required to package the information in correspondence with the most current common ground.

Information can be introduced to the common ground either in terms of propositions (Stalnaker 1974; Karttunen 1974; Lewis 1979) or in terms of discourse referents (Kamp 1981; Heim 1982). Propositions are concerned with the truth-conditions of the information which are mutually accepted by the interlocutors and are associated with the *common ground content*. Discourse referents, on the other hand, are concerned with the development of the common ground, which falls under the notion of *common ground management* (Clark 1996; Groenendijk 1999; Merin 1994). Discourse referents can be explicitly introduced, accommodated or taken up by different lexical expressions. The different linguistic expressions signal information about communicative interests and goals of the interlocutors.

2.1.1 Givenness

An essential part of information structure is the notion of *givenness*. Its inflationary use, however, resulted in multiple definitions being subsumed under the single notion of givenness. For instance, Prince (1981) discerns three different terms of givenness, namely $Givenness_P$, $Givenness_S$ and $Givenness_K$, which all differ to some extent in their meaning.[2] Thus, providing a precise definition is crucial in order to establish a common ground with the reader.

Krifka & Musan (2012) have contributed a rather theory neutral description of givenness which applies to both the lexical entry level as well as the constituent level. They define givenness as a feature of an expression α which "indicates whether the denotation of α is present in the common ground or not, and/or indicates the degree to which it is present... in the immediate common ground" (Krifka & Musan 2012: 22). This thesis will not be concerned with givenness as a feature of lexical items, which includes, for example, personal pronouns, but as a marker of constituents, especially those which denote discourse referents.

[2]The subscripted letter P stands for *predictability*, S for *salience* and K for *shared knowledge* (Prince 1981).

Moreover, it will be assumed that givenness does not constitute a binary feature. Instead, it will be regarded to be gradient in nature, thereby resembling a continuous spectrum containing various degrees of givenness levels. This is in line with Prince (1981), who describes her degrees of givenness under the heading of *assumed familiarity*, and with Chafe (1994; 1996), who accounts for his degrees of givenness in terms of *activation cost*. In the following, I will go into more detail on Prince's (1981) and Chafe's (1994) notions of givenness.

Prince's (1981) assumed familiarity is based on the assumption that a speaker and a listener share knowledge about what they assume to be in the consciousness of the other person. The different linguistic expressions used by the speaker in order to refer to a discourse referent reveal the kinds of assumptions the speaker makes about what is part of the listener's consciousness. The listener, on the other hand, draws inferences about the discourse referent on the basis of the lexical expression used by the speaker.

Prince (1981) distinguishes between three main levels of givenness: *Evoked, new* and *inferable*. As depicted in Figure 1, each level is subdivided into further categories. What follows is a description of each givenness sublevel.

EVOKED. A referent is considered to be *textually evoked* if it has been explicitly mentioned by the speaker. Textually evoked referents can be introduced to the common ground either as new or inferred referents. A referent is considered to be *situationally evoked* if it is selected from the immediate surroundings of the interlocutors. Thus, situationally evoked referents include the information about the interlocutors as well as the extra-textual context.

NEW. A referent is considered to be *unused* if the speaker assumes that the listener has knowledge about the referent but is not conscious of it as it has not been introduced to the common ground, yet. This often involves prominent figures like *Angela Merkel* or *The Queen*. A referent is considered to be *brand-new unanchored* if the speaker assumes that the listener has no prior knowledge about the referent. The unanchored brand-new referent is often referred to with an indefinite noun phrase, such as *a friend*. A referent is considered to be *brand-new anchored* if the speaker introduces an unknown referent in relation to a known one by using a noun phrase which contains another noun phrase serving as a link for the unknown referent. For instance, in the indefinite noun phrase *a friend of mine*, the personal pronoun *mine* serves as an anchor for the brand-new referent *a friend*. It is crucial that the anchor must not be brand-new.

INFERABLE. A referent is considered to be *noncontaining inferable* if the referent can be inferred from an already evoked or inferable discourse referent

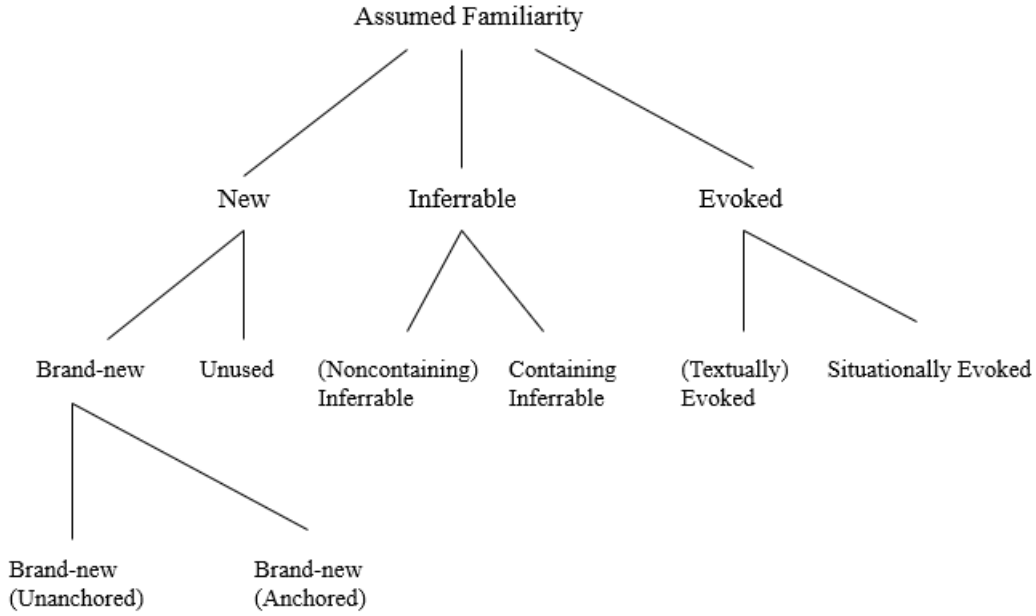

Figure 1: Degrees of givenness by Prince (1981: 73).

by means of "logical or... plausible reasoning" (Prince 1981: 236). This is exemplified in (3) (Prince 1981: 233), where *the driver* can be "plausibly" inferred from *a bus*. A referent is considered to be *containing inferable* if the referent is inferred from a set it is a part of. Thus, in *one of my friends*, it is not only the referent *one* but the whole partitive construction which is considered to be the containing inferable. All inferable referents are definite and cannot, for instance, be marked with an indefinite article.

(3) I got on a bus yesterday and the driver was drunk.

In line with Prince (1981), Chafe (1994; 1996) also distinguishes between three levels of givenness.[3] He refers to them as *given, new* and *accessible*. Chafe (1994; 1996) defines the three levels of givenness as the mental cost a listener needs to invest in order to transfer a referent from a prior activation state to an active one, hence the term activation cost. As can be seen in Figure 2, given referents do not undergo a change in activation state as they are already active, accessible referents change from a semi-active state to an active state and new referents change from an inactive state to an active state. The amount of activation cost that is imposed on the listener increases from given referents being the least costly, accessible referents being more costly and new referents being the most costly, "presumably because more mental effort is involved in converting an idea from the inactive to the active state" (Chafe 1994: 73).

[3]Chafe (1994) does not rule out the existence of more than three givenness levels.

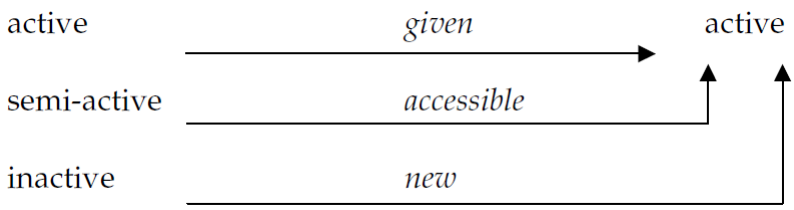

Figure 2: Degrees of givenness by Chafe (1994: 73).

Chafe (1994) describes the three levels of givenness in relation to the speaker's assumption about the listener's state of consciousness at the time of the discourse utterance. Thus, the speaker assumes that given referents are part of the active consciousness of the listener while new referents are assumed to be newly introduced to the listener's consciousness by what the speaker says. For a referent to newly enter the listener's consciousness, it can either be assumed to be previously unknown or already known to the listener. The referent is considered to be new as long as the listener is made actively conscious of it at the time of the utterance.

Accessible referents are assumed to be in the "peripheral consciousness" (Chafe 1996: 40) of the listener. They enter the semi-active state via inference from a given context (Chafe 1996).[4] The listener must identify the accessible referent as an essential part of an idea that is expressed in the discourse. Put differently, the accessible referent must be "necessarily included in the mental image of what is being talked about" (Chafe 1996: 44). In example (3), it is the mental image of a bus which necessarily includes the mental image of a driver, thereby making *the driver* referentially accessible from *a bus*.

A speaker will choose his or her linguistic expressions based on the assumed activation state of the referent. Thus, if the speaker assumes that a referent is already active in the listener's consciousness, he or she will often chose "weakly accented pronouns" (Chafe 1994: 75) to refer to the given referent. If the speaker assumes that a referent is either semi-active or inactive in the listener's consciousness, he or she will chose "accented full noun phrases" (Chafe 1994: 75). What is noticeable is that although Chafe (1994; 1996) ascribes two different functions to semi-active and inactive states, he claims that they do not differ with regard to their linguistic form. However, studies conducted on German have shown that each activation state, indeed, has a distinct prosodic correlate. While given referents are preferably deaccented, new referents tend to be realised with rising pitch accents and accessible referents tend to be realised with falling pitch accents (Baumann 2006; Baumann 2008; Baumann

[4]Another possibility for a referent to enter the semi-active state is by receding from a previously active state (Chafe 1994; 1996).

& Grice 2006; Baumann et al. 2015). The relation between givenness and prosody will be described in detail in §2.4.

2.1.2 Focus

There have been considerable differences in the way the notion of *focus* have been used in the literature. Chomsky (1969) and Jackendoff (1972) argue that *discourse-new* and *contrastive* referents belong to the same focus category because they do not differ with regard to their phonological representation. According to them, there is no phonological difference between the answers in (4), which identifies *Mary* as discourse-new, and in (5), which identifies *Mary* as contrastive.[5]

(4) A: What happened? / Who did John kiss?
B: John kissed [MAry.]$_{\text{NEW}}$

(5) A: Did John kiss Susan?
B: John kissed [MAry.]$_{\text{FOCUS}}$

Katz & Selkirk (2011), however, provide prosodic evidence against the claim by Chomsky (1970) and Jackendoff (1972). They measured the phonetic realisation of English minimal sentence pairs which differed only with regard to a single constituent being either newly or contrastively marked. They found that both newly and contrastively marked constituents had similar pitch accentuation and phonological phrase properties. However, contrastively marked constituents were found to exhibit longer durations, greater relative intensities and greater fundamental frequency movements, thereby being more prosodically prominent than their discourse-new counterparts. Katz & Selkirk (2011) argue that this difference in prosodic prominence cannot be explained in terms of differences in pitch accent types or phrasal boundaries. Instead, they attribute this difference to the distinct grammatical status of contrastive focus.

Selkirk (1995) and Chafe (1976; 1994), among others, also distinguish between new and contrastive interpretations of focus acknowledging only the latter. According to them, focus indicates "the presence of alternatives that are relevant for the interpretation of linguistic expressions" (Krifka & Musan 2012: 7). Chafe (1976) believes that the "speaker assumes that a limited number of candidates is available in the addressee's mind" (Chafe 1976: 34). The answer in (5), then, is chosen from a set of alternatives which contains sentences that are identical to one another except for the subconstituent in focus. An alternative set of (5) is exemplified in (6).

[5]Capital letters indicate pitch accentuation.

(6) {John kissed Mary, John kissed Susan, John kissed Helen, ...}

According to Chafe (1976), it is necessary that the size of the alternative set is limited. This proposition is based on Bolinger (1961) who states that "in a broad sense every semantic peak is contrastive" and "[a]s the alternatives are narrowed down, we get closer to what we think of as a contrastive accent" (Bolinger 1961: 81). Contrastiveness and discourse-newness can, therefore, be thought of as part of a continuum where one extreme signals a single alternative while the other signals an unlimited set of alternatives. The smaller the set of alternatives the higher the probability that the sentence is considered to be contrastive. In contrast, an unlimited set of alternatives, which, for example, can be evoked by the general question *What happened?* fails to be contrastive. Thus, example (4) can be associated with an unlimited set of alternatives.

Concerning the nature of alternative sets, Krifka & Musan (2012) distinguish between *expression focus* and *denotation focus*. Expression focus contrasts constituents either on the morphological or on the phonological level. As shown in (7), the constituents in focus are not contrastive in terms of their meaning.

(7) Expression focus

a. Sie ist nicht [FETT]$_{\text{FOCUS}}$ sondern [FLAUschig.]$_{\text{FOCUS}}$

b. A: Die Leichtathletik-WM war diesmal in [KAtar].
B: Sie war in [KaTAR!]$_{\text{FOCUS}}$.

The alternatives of both foci in (7a) are part of the set {fett, flauschig}. Albeit differing with regard to their connotations, both expressions cannot have their own denotation as they refer to the same property of being heavy. The alternatives in (7b) are part of the set {KAtar, KaTAR}. Both expressions are contrastive in terms of prosody but not in terms of meaning. It can be assumed that speaker B corrects the pronunciation of speaker A without knowing that both pronunciations are correct in German (Dudenredaktion n.d.).

Unlike expression focus, denotation focus does carry meaning on its own. According to Krifka & Musan (2012), "[t]he relevant alternatives are construed on the level of denotations, leading to alternative denotations of complex expressions" (Krifka & Musan 2012: 8). Importantly, the alternative denotations have to be similar to the denotation of the expression in focus with regard to their grammatical category as well as their ontological being. Thus, if the denotation in focus is a human-being which is referred to by means of a definite noun phrase, all denotations in the alternative set must be human-beings which are referred to with a definite noun phrase as well. This thesis will be concerned with denotation focus as opposed to expression focus.

2.1.3 Topic

Another highly ambiguous term is the notion of *topic*. Its most frequent usage has been that "one part of a sentence says something about another part" (Krifka & Musan 2012: 25). According to this definition, topics refer to the part of the sentence the sentence is about. The part of the sentence which says something about the topic is referred to as *comment*. Partitioning a sentence into topic and comment can be ascribed to Hockett (1958). Unlike Hockett (1958) who attributes this distinction to the syntactic structure of the sentence alone, Reinhart (1981) attributes this distinction to linguistic as well as information structural factors.

According to Reinhart (1981), the storage of information in the common ground can be understood in terms of a file card system which was used by libraries in a time before the digitisation. The information the speaker identifies as the topic can be considered as the heading of the file card while information identified as the comment about the topic are filed under that heading. Within this model, "[t]he topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the common ground content" (Krifka & Musan 2012: 28). Thus, the information in the common ground is structured in such a way that discourse referents are paired with propositions about these discourse referents. This allows for sentences that express the same propositions to have different structures. As exemplified in (8), (8a) is stored under the heading *John* while (8b) is stored under the heading *Mary*.

(8) a. [John]$_{\text{TOPIC}}$ [kissed Mary.]$_{\text{COMMENT}}$
b. [Mary]$_{\text{TOPIC}}$ [kissed John.]$_{\text{COMMENT}}$

What is also exemplified in (8) is that the topic is located at the left-most position of the sentence. Defining the topic as the first element of the sentence can be traced back to the Prague School and was later adopted by Halliday (1967; 1970), among others. According to Halliday, the left periphery of a sentence is reserved for the topic and functions as "the point of departure for the clause as a message" (Halliday 1967: 212) and the "the peg on which the sentence is hung" (Halliday 1970: 161).[6] Crucially, the topic of the sentence is neither required to be the grammatical subject of the sentence nor to be information structurally given. It is merely "what comes first in the clause" (Halliday 1967: 212) which counts as topical. Halliday (1967), therefore, identifies *John*, *yesterday* and *the play* as the topics in (9) (Halliday 1967: 212).

[6]Halliday (1967; 1970) uses the notion of *theme*.

(9) a. [John]$_{\text{TOPIC}}$ saw the play yesterday.
b. [Yesterday]$_{\text{TOPIC}}$ John saw the play.
c. [The play]$_{\text{TOPIC}}$ John saw yesterday.

In contrast, Givón (1983) argues that only discourse referents have the potential of becoming the topic of a sentence. Discourse referents have the ability to form "thematic paragraphs" (Givón 1983: 7) by keeping their topic status in the following sentences as well. Thus, a discourse referent which is the center of a "chain of equi-topic clauses" (Givón 1983: 9) can be characterised as highly continuous by forming the basis for the aboutness of the thematic paragraph. According to (Givón 1983), then, the topics in (9) would be *John* and *the play* as shown in (10).

(10) a. [John]$_{\text{TOPIC}}$ saw the play yesterday.
b. Yesterday [John]$_{\text{TOPIC}}$ saw the play.
c. [The play]$_{\text{TOPIC}}$ John saw yesterday.

2.2 Centering Theory

Centering Theory, developed by Grosz et al. (1983; 1995), is considered to be one of the most influential theories for reference tracking in discourse. It models how the choice of referential expression of the speaker and the inference load on the part of the listener contribute to the coherence of the utterances within a discourse. A set of rules governs the coherence of a discourse segment. As the name of the theory implies, these rules are referred to as *centering rules*, which revolve around the so-called *center*.

A discourse segment consists of a sequence of utterances. A sequence of utterances differs from a sequence of solitary sentences in that it contains referents which connect the utterances with each other, thereby creating a coherent discourse. Within the framework of Centering Theory, referents are called *centers*. Centers do not constitute linguistic expressions but semantic objects which are denoted by noun phrases. The denotation or linguistic realisation of a center within a discourse segment combines levels of syntax, semantics, pragmatics and cognition.

The fundamental idea of Centering Theory is that "certain referents of an utterance are more central than others" (Schumacher et al. 2015: 11). As exemplified in (11) and (12) (Grosz et al. 1995: 206), a speaker's choice of referential expression combined with the referent's grammatical role reflects the different degrees of centrality of the referent, which further affects the degree of coherence within the discourse segment.

(11) a. John went to his favourite music store to buy a piano.
b. He had frequented the store for many years.
c. He was excited that he could finally buy a piano.
d. He arrived just as the store was closing for the day.

(12) a. John went to his favourite music store to buy a piano.
b. It was a store John had frequented for many years.
c. He was excited that he could finally buy a piano.
d. It was closing just as John arrived.

Example (11) and (12) show two discourse segments. Contentwise both discourse segments are identical, yet discourse (11) seems to exhibit more coherence than discourse (12). In (11), the referent denoted by the proper name *John* is the sole center of attention. *John* remains in subject position and is referred to by a pronoun in all subsequent utterances. (12), on the other hand, lacks a clear center of attention. Rather, the aboutness of the discourse alternates between the referent denoted by *John* and another referent denoted by the noun phrase *his favourite music store*. In addition, the referential expression used to refer to *John* alternates between a pronoun in utterance-initial position and a proper name in utterance-medial position. These examples show that linguistic expressions combined with the grammatical role directly impact a referent's degree of centrality as well as the overall discourse coherence.

Discourse utterances are connected via centers. Each utterance U is assigned a set of forward-looking centers ($C_f(U)$). In contrast, each utterance U other than the segment-initial utterance is only assigned a single backward-looking center ($C_b(U)$). The backward-looking center of an utterance connects with one of the forward-looking centers of the previous utterance. Every element in the set of the forward-looking centers in one utterance ($C_f(U_n)$) has the ability to become the next backward-looking center in the subsequent utterance ($C_b(U_{n+1})$). The likelihood of an element of $C_f(U_n)$'s to become the next $C_b(U_{n+1})$ is determined by the relative discourse prominence of all elements within that set. The more discourse prominent an element is, the higher it scores in the ranking. Hence, the element that is ranked highest becomes the backward-looking center in the subsequent utterance. Crucially, $C_b(U_n)$ can only be chosen from the set of the forward-looking centers $C_f(U_{n-1})$ of the previous utterance, not from prior sets, such as $C_f(U_{n-2})$. It is, therefore, strictly local.

The way elements within the set of forward-looking centers are chosen is two-fold. First, they are chosen based on the utterance in which they occur. In other words, the $C_f(U_n)$'s are dependent on the referential expressions that

occur in U_n. Second, forward-looking centers can be chosen based on "utterances that contain noun phrases that express functioning relations" (Grosz et al. 1995: 217).[7] Instead of being directly expressed in U_n, the $C_f(U_n)$'s are indirectly inferred from the $C_b(U_{n-1})$ that is denoted by a noun phrase expressing functioning relations. Consider discourse (13) which is about a house that is denoted by the noun phrase *the house* in (13a). *The door* in (13b) and *the furniture* in (13c) are functionally dependent on *the house* in that they constitute parts of the house. Because of *the house* being the sole referent in (13a), it becomes the $C_b(U_1)$ in (13b). The set of $C_f(U_2)$'s contains the newly introduced referent *the door* but also *the house*, which is being realised indirectly via *the door* (Grosz et al. 1995: 217). *The door* is said to outrank *the house* because U_2 revolves around *the door*. Interestingly, *the house* continues to be the $C_b(U_2)$ in the subsequent utterance, thereby violating the ranking of $C_f(U_2)$'s. Aboutness as well as processing similarities between U_1 and U_2 seem to factor in the choice of backward-looking centers when noun phrases with functioning relations are involved.

(13) a. [The house appeared to have been burgled.]$_{U_1}$
$C_f(U_1) = \{$the house$\}$

b. [The door was ajar.]$_{U_2}$
$C_b(U_1)$ = the house; $C_f(U_2) = \{$the door, the house$\}$

c. [The furniture was in disarray.]$_{U_3}$
$C_b(U_2)$ = the house; $C_f(U_3) = \{$the furniture, the house$\}$

There are three ways in which an utterance transitions to the next utterance. They transition by either continuing, retaining or shifting their center. Consider example (14) by Grosz et al. (1995: 217). (14a) introduces a referent denoted by the proper name *John*. In (14b), the pronoun *he* can only connect back to *John* as he is the sole backward-looking center. Consequently, he is also the sole referent contained in the set of forward-looking centers. Being without competition, *John* CONTINUES to be the backward-looking center in (14c). Due to the newly introduced referent denoted by *Harry*, the set of forward-looking centers, now, comprises both, *John* and *Harry*. As *John* outranks *Harry* in (14c), the pronoun *him* in (14d) connects back to *John*. Due to syntactic and lexical factors, *Harry*, now, outranks *John* in (14d). Thus, *John* cannot become the next backward-looking center in (14e); he is merely RE-

[7] Clark (1977) uses the term *bridging* relations.

TAINED. In the end, the backward-looking center of (14e) SHIFTS from being *John* in (14d) to being *Harry* in (14e).

(14) a. [John$_{\text{i}}$ has been having a lot of trouble arranging his vacation.]$_{\text{U}_1}$

b. [He$_{\text{i}}$ cannot find anyone to take over his responsibilities.]$_{\text{U}_2}$
$\text{C}_\text{b}(\text{U}_1)$ = John; $\text{C}_\text{f}(\text{U}_2)$ = {John}

c. [He$_{\text{i}}$ called up Harry$_{\text{j}}$ yesterday to work out a plan.]$_{\text{U}_3}$
$\text{C}_\text{b}(\text{U}_2)$ = John; $\text{C}_\text{f}(\text{U}_3)$ = {John, Harry} (CONTINUE)

d. [Harry$_{\text{j}}$ has annoyed him$_{\text{i}}$ a lot recently.]$_{\text{U}_4}$
$\text{C}_\text{b}(\text{U}_3)$ = John; $\text{C}_\text{f}(\text{U}_4)$ = {Harry, John} (RETAIN)

e. [He$_{\text{j}}$ called John$_{\text{i}}$ at 5 AM on Friday last week.]$_{\text{U}_5}$
$\text{C}_\text{b}(\text{U}_4)$ = Harry; $\text{C}_\text{f}(\text{U}_5)$ = {Harry, John} (SHIFT)

Shifting centers across utterances, such as in (14e) from *John* to *Harry*, requires additional inference load on the part of the listener. In comparison, continuing with the same center, as it is in (14c), requires less inference load. As exemplified by the alternating choice of *John* and *he* in (12) above, the chosen referential expression can affect the amount of inference load placed upon the listener as well. Inference load and discourse coherence are intrinsically linked with one another in that inference load is reversely proportional to a listener's perceived coherence. Thus, a discourse segment becomes all the more coherent the smaller the inference load is.

Centering rules are based on the assumption that there is a preference for discourse structures to be more coherent, therefore requiring less inference load. Hence, adhering to those rules will lead to a more coherent discourse segment while violating them will lead to a less coherent one. The determining factors according to which forward-looking centers are ranked are grammatical role, choice of referential expression and type of transition across utterances. Forward-looking centers that are denoted by subjects rank higher than those denoted by objects, which in turn rank higher than other grammatical roles. Grosz et al. (1995: 214) establish the following ranking for grammatical role:

SUBJECT > OBJECT(S) > OTHER

Moreover, the choice of referential expression for referents of lower ranked grammatical roles is being constrained in that they cannot be pronominalised unless higher ones are. This is established by Rule 1 which states that "no element in an utterance can be realised as a pronoun unless the backward-looking

center is realised as a pronoun also" (Grosz et al. 1995: 214). Importantly, Rule 1 can only apply given that the highest ranking forward-looking center of an utterance U_n is actually uttered with a pronoun in U_{n+1}. The pronominalised (or given) backward-looking center is, then, more likely to be picked up again as the next backward-looking center in U_{n+2}. Note that the inference load will increase if the backward-looking center in U_{n+2} will be realised by a noun phrase instead of a pronoun, as exemplified in (12c) and (12d) above. Thus, inference load can increase without necessarily violating Rule 1.

> Rule 1
> If any element of $C_f(U_n)$ is realised by a pronoun in U_{n+1},
> then the $C_b(U_{n+1})$ must be realised by a pronoun also.

> Rule 2
> $\textsc{continue}(U_n,U_{n+1}) > \textsc{retain}(U_n,U_{n+1}) > \textsc{shift}(U_n,U_{n+1})$

Another rule establishing discourse coherence is concerned with the movement of centers. Rule 2 states that "sequences of continuations are preferred over sequences of retaining; and sequences of retaining are preferred over sequences of shifting" (Grosz et al. 1995: 215). Moreover, the order in which transition types occur can affect inference load in different ways. Continuations that are immediately followed by shifts require higher inference load than those which are initially followed by retentions, which are, then, followed by shifts. Retentions facilitate processing by smoothing out the transition process.

2.3 Discourse prominence

Within the framework of Centering Theory, Grosz et al. (1995) claimed that for a referent to become the next backward-looking center it must be more discourse prominent than other referents. The notion of discourse prominence can be accounted for by the *givenness hierarchy* proposed by Gundel et al. (1993), the *topic accessibility* by Givón (1983) and the *accessibility marking scale* by Ariel (1988; 1990; 1991; 1996).[8] All hierarchies suggest that the degree of discourse prominence is closely linked to the linguistic expression used for a referent.

According to Gundel et al. (1993), a referent's discourse prominence can be attributed to the referent's degree of givenness in the ongoing discourse. They

[8]The notion of accessibility used by Givón (1983) and Ariel (1991) should not be confused with the one used by Chafe (1994; 1996).

propose a hierarchy of discourse prominence which employs the notions of givenness by Chafe (1994) and Prince (1981). Referents which rank higher in the givenness hierarchy are considered to be more activated whereas referents which rank lower in the givenness hierarchy are considered to be less activated. More activation coincides with more discourse prominence while less activation coincides with less discourse prominence.

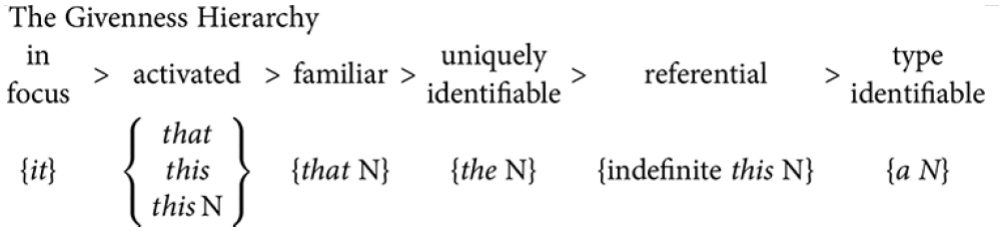

Figure 3: Hierarchy of discourse prominence by Gundel et al. (1993: 275).

As can be seen in Figure 3, the givenness hierarchy consists of six levels which are closely associated with specific referential expressions. Unlike Prince's (1981) notion of givenness, the levels of givenness by Gundel et al. (1993) bear entailment relations, such that higher givenness levels contain lower ones, but not vice versa. Thus, each level in the hierarchy constitutes a necessary and sufficient condition for the appropriate use of a referential expression. Using a specific referential form signals that the associated givenness level as well as all lower levels located on the right hold. Furthermore, referents *in focus*,[9] which rank highest in the givenness hierarchy, are identified from the most restrictive set of referents while *type identifiable* referents, which rank lowest in the givenness hierarchy, are identified from the least restrictive set of referents.

As for the referential form, it is not the case that the referential expressions of less given referents are necessarily available for more given ones. Consider example (15) in which the referent *John* is in focus.

(15) a. John$_i$ went to the store.
b. He$_i$ bought a piano.
c. *A man$_i$ left the store.

Referring back to *John* with the indefinite noun phrase *a man* is regarded to be inappropriate. Gundel et al. (1993) argue that speakers take into account Grice's (1975) *Maxim of Quantity* which postulates for the current purpose of exchange that a speaker is required to make their contribution as informative as required (Q1), but not more informative than required (Q2). Thus, the six

[9]Gundel et al.'s (1993) notion of focus should not be confused with the one by Chafe (1976; 1994) and Selkirk (1995).

levels of givenness form an implicational scale constraining the availability of referential forms. The use of the infinite noun phrase in (15c) maximally violates Q1 as it informs the reader that *John* is only a type identifiable referent when, in fact, he is a referent in focus.

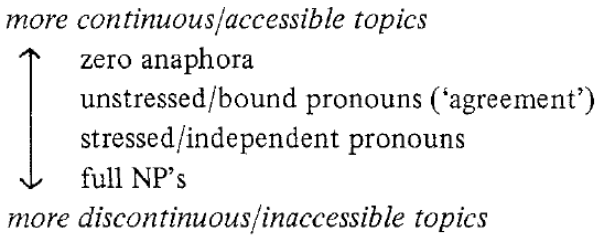

Figure 4: Hierarchy of discourse prominence by Givón (1983: 18).

Another hierarchy of discourse prominence have been proposed by Givón (1983). He claims that the accessibility of a discourse referent (or topic) is dependent on four factors: The referential distance, the potential interference by alternative referents, the availability of semantic information as well as the availability of thematic information. All four factors interfere with the process of identifying the topic and hamper the production of a "chain of equi-topic clauses" (Givón 1983: 9). Givón (1983) argues that topics with higher accessibility are more continuous while those with lower accessibility are less continuous. This is illustrated in Figure 4.

Figure 4 shows that the degree of a referent's discourse prominence has direct implications for the phonological size of the referential expression. That is "[t]he more disruptive, surprising, discontinuous or hard to process a topic is, the more coding material must be assigned to it" (Givón 1983: 18). Thus, reduced forms such as zero anaphoras or "unstressed" pronouns have higher degrees of accessibility and continuity, which is why they are considered to be more discourse prominent than "stressed" pronouns or full nouns phrases which are considered to be less accessible and rather discontinuous.

Lastly, Ariel (1988; 1990; 1991; 1996) proposes, as part of her *Accessibility Theory*, a hierarchy of discourse prominence which assumes that a referent's degree of accessibility is coded in the linguistic expression. Thus, a speaker's choice of linguistic expression takes into consideration to what degree he or she assumes that the listener has access to the referent. After having established the referent's degree of accessibility, the speaker chooses the appropriate linguistic expression. As can be seen in Figure 5, the linguistic expressions are arranged hierarchically on a scale of accessibility. The listener, then, searches

in his or her mental representation for a referent that suits the degree of accessibility which have been expressed by the chosen linguistic form best.

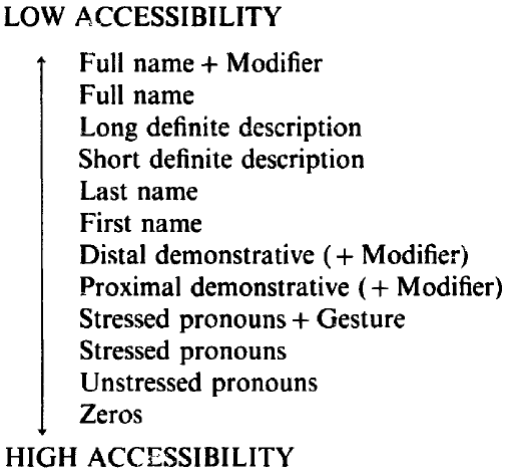

Figure 5: Hierarchy of discourse prominence by Ariel (1991: 449).

The discourse hierarchy proposed by Ariel (1991) resembles the one by Givón (1983) in that more linguistic coding signals lower accessibility of the referent while less linguistic coding signals higher accessibility of the referent. The discourse hierarchy is based on three criteria: Informativity, rigidity and attenuation of the linguistic expression. Informativity indicates the lexical and semantic richness of a referential expression. For instance, pronouns like *sie* carry less semantic information for which they are considered to be less informative than definite descriptions like *die Studentin*. Rigidity indicates "how uniquely referring an expression is" (Ariel 1991: 450). For instance, proper names like *Klaus von Heusinger* are considered to be more uniquely referring than definite descriptions like *der Germanistikprofessor*, considering that there are more than one professors of German language at the University of Cologne, let alone in the world. Attenuation distinguishes between linguistic expressions which are equally informative but differ in terms of their phonological size. For instance, an "unstressed" pronoun is more attenuated than its "stressed" counterpart. Taken together, "[t]he more informative, the more rigid and the least attenuated the form the lower Accessibility it marks, and vice versa" (Ariel 1991: 449).

What catches the eye of a phonologist immediately is that the meaning of "unstressed" and "stressed" is not elaborated on further. The following section will address the role of prosodic prominence in more detail.

2.4 Prosodic prominence

One of the central functions of prosody is to highlight information by making some elements more prosodically prominent than others. Gussenhoven (2004) proposes that speakers follow the so-called *Effort Code*, according to which highlighting information requires more articulatory effort, which in turn, results in wider excursions of pitch movements. The relation between highlighted information, prosodic prominence and pitch movement can, therefore, be regarded as linear in that greater pitch movements correlate with greater prosodic prominence, which in turn indicates greater informativity.

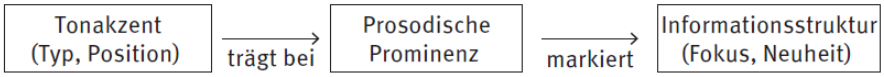

Figure 6: Relation between pitch accent, prosodic prominence and information structure in German (Grice & Baumann 2016: 102).

Figure 6 (Grice & Baumann 2016: 102) shows a schematic representation of the role of prosodic prominence acting as a connecting element between pitch accentuation, on the one hand, and information structure, on the other hand. In the following, I will go into more detail about the different German pitch accent types and how they relate to the information structural notions of givenness and focus. The prosodic structure of German will be addressed using the *German Tones and Breaks Indices* or in short *GToBI* (Grice et al. 1996; Reyelt et al. 1996).

2.4.1 GToBI

Talking about prosody requires a framework phonologists consensually agree upon. One such framework is GToBI (Grice et al. 1996; Reyelt et al. 1996). GToBI is a phonological transcription tool which was specifically developed for the intonation of German. It is based on the original *ToBI* framework developed for American English by Beckman & Hirschberg (1994) and Beckman et al. (2005). One of the basic assumptions of GToBI is that utterances, or rather intonation units, consist of a *text* level, which contains the segmental components of an utterance, and a *tune* level, which contains the suprasegmental components of an utterance. This is in the spirit of the *Autosegmental-Metrical Phonology* proposed by Ladd (2008) who uses the term *autosegments* to refer to the levels of text and tune. According to Ladd (2008), both levels behave autonomously from one another but are connected via various phonological processes.

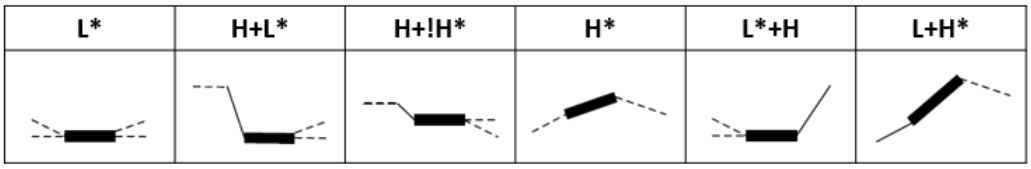

Figure 7: Schematic representation of German pitch accent types proposed by GToBI. Bold lines indicate the direction of tonal movements on the pitch accented syllable. Solid lines before and after the pitch accented syllable indicate obligatory tonal movements and dashed lines indicate facultative tonal movements (Grice & Baumann 2000: 12).

GToBI differentiates between six pitch accent types which consist of a combination of high (H) and low (L) tones. Its inventory contains two monotonal (H* and L*) and four bitonal (L*+H, L+H*, H+L* and H+!H*) pitch accents.[10] Their schematic representations are shown in Figure 7. The different pitch accent types describe the perceived tonal contour of an intonation unit. Starred tones are generally associated with metrically strong syllables resulting in the highlighting of the constituent the syllable is a part of. As for bitones, the plus sign symbolises that the tonal sequence forms a unit. Moreover, preceding unstarred tones are referred to as *leading tones* while following ones are referred to as *trailing tones*. Leading and trailing tones indicate target points in the close vicinity of the pitch accented syllable. The diacritic '!' indicates that the corresponding pitch accent is affected by *downstep*, a step-wise lowering of the pitch. All pitch accents differ with regard to three tonal dimensions. First, the direction of tonal movement which is either *rising* (e.g. L+H*) or *falling* (e.g. H+L*). Second, the association of *pitch peaks* (e.g. H*) or *pitch valleys* (e.g. L*) with metrically strong syllables. Lastly, the scaling and height of pitch. Crucially, pitch accents should not be understood in absolute terms but relative to a speaker's pitch range. As a rule of thumb, high tones are considered to be located in the upper quarter while low tones are located in the lower quarter of a speaker's pitch range.

The relative perceived prosodic prominence for the different tonal categories of GToBI including deaccentuation have been examined in a perception experiment by Baumann & Röhr (2015). The aim of the study was to directly compare the prosodic prominence of tonal categories and not the contextual appropriateness of a specific pitch accent or lack thereof. Native speakers of German were asked to evaluate the degree of perceived 'highlightedness' or *Hervorgehobenheit* of contextless sentences which were identical apart from

[10]Over time, this number varied between five (Grice & Baumann 2002; 2016) and six (Grice et al. 2005).

their intonation contours. The target word was either deaccented or realised with one of the following pitch accent types L*, H+L*, H+!H*, !H*, H*, L*+H or L+H* in nuclear position.

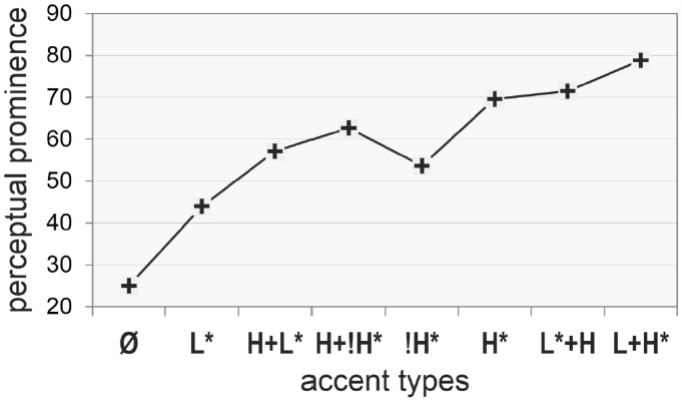

Figure 8: Mean evaluations of the perceived prosodic prominence of different German pitch accent types including deaccentuation (Baumann & Röhr 2015: 4).

Baumann & Röhr (2015) found gradual differences between all pitch accent types including deaccentuation. The results are shown in Figure 8. Deaccentuation was found to be the least prominent tonal category. Differences in the prominence evaluations of the pitch accent types can be summarised in terms of three tonal dimensions: The direction of pitch movement, the degree of pitch excursion and the height of the starred tone. As for the direction of pitch movement, rises were found to be perceived more prominent than falls (e.g. L+H* > H+L*). Moreover, bitonal pitch accents, which tend to have steeper pitch excursions, were found to be more prominent than monotonal pitch accents, which have more shallow pitch excursions (e.g. L+H* > H*). Lastly, high pitch accents were found to be more prominent than downstepped pitch accents, which in turn were more prominent than low ones (e.g. H* > H+!H* > L*). This study showed that native speakers of German are able to differentiate not only between the presence and absence of pitch accents but also between different types of pitch accents.

2.4.2 Prosody of focus

Every intonation unit has at least one pitch accent which is referred to as the *nucleus* or the *nuclear pitch accent* (Grice & Baumann 2016). The nucleus is structurally the most prominent pitch accent and is typically placed on the last argument of the phrase, unless this argument is pronominalised (Uhmann 1991; Ladd 2008). Other pitch accents which belong to the same intonation unit are described in relation to the nucleus. For example, pitch accents which occur to the right of the nucleus are referred to as *prenuclear pitch accents*.

Prenuclear pitch accents are perceived to be less prominent than their nuclear counterparts (Jagdfeld & Baumann 2011).

In German, the information structure can override the default pitch accent placement of the nucleus. That is, the type and position of the nucleus pitch accent is determined by the focus-background structure of the utterance. Focus denotes the part of an utterance which the speaker assumes to be most informative for the listener (Lambrecht 1994). Background, on the other hand, denotes the uninformative part of an utterance. In a speech production study, Mücke & Grice (2014) investigated how the important part of an utterance is differentiated from the uninformative part. They tested four types of focus-background structures which are exemplified in Figure 9.

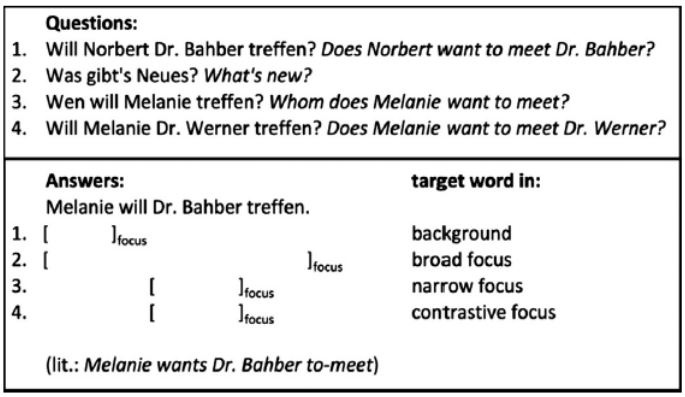

Figure 9: Different focus-background structures used in the production study by Mücke & Grice (2014: 52).

In the background condition, the target word *Dr. Bahber* was already given in the context and therefore deaccented. Hence, the nuclear pitch accent was on the first argument of the utterance. In the broad focus condition, the whole utterance was focused because of which the nuclear pitch accent was, by default, on the last argument. The narrow and contrastive focus conditions signalled the presence of alternatives to the target word. While alternatives in the narrow focus condition were marked implicitly, those in the contrastive focus condition were mentioned explicitly in the immediate context. In both cases, the nuclear pitch accent was on the last argument.

Mücke & Grice (2014) found that the target word in the background condition did not receive a nuclear pitch accent. The target word in the other three focus conditions, however, did. As can be seen in Figure 10, different types of nuclear pitch accents were preferred for different focus structures. H* was realised in all three focus conditions and can, therefore, be regarded to be pragmatically neutral. H+!H* was preferred in the broad focus but was never used to mark contrastive focus. L+H* was predominantly used to mark con-

trastive focus but was often used to mark narrow focus as well. Thus, L+H* can be regarded as a marker for implicit and explicit alternatives.

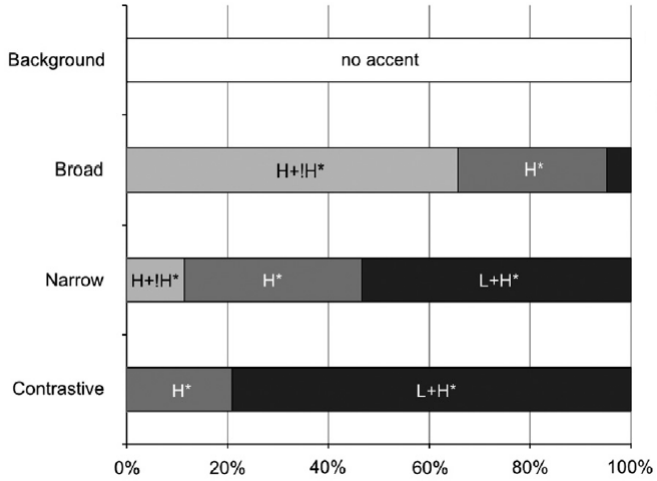

Figure 10: Distribution of pitch accents on the target word (e.g. Dr. Bahber) in different focus conditions (Mücke & Grice 2014: 53).

The question whether it is the pitch accent type or the focus domain that contributes to the activation of alternatives was investigated in two visual world eye-tracking experiments by Braun (2015). Braun (2015) compared the processing of L+H* and H+L* on narrowly focused subject arguments to the processing of L+H* on broadly focused subject arguments in phrase-initial position. An example item is given in (16). While listening to the items, participants were visually presented with four competing referents: The non-contrastive associate (e.g. *Turner*), a contrastive associate (e.g. *Tänzer*), the grammatical object (e.g. *Blasen*) and an unrelated distractor.

(16) TEST: [Der Turner]$_{\text{FOCUS}}$ hatte Blasen bekommen.
L+H* ∅
H+L* ∅
CONTROL: [Der Turner hatte Blasen bekommen.]$_{\text{FOCUS}}$
L+H* L+H*

Braun (2015) found significantly longer eye fixations on the contrastive associate when the subject was produced with L+H* in narrow focus than in broad focus. No such effect was observed for H+L*. These findings suggest that the activation of alternatives is dependent not only on the type of focus domain but also on the type of pitch accent: While narrowly focused constituents being realised with a nuclear L+H* activate alternatives, nuclear H+L* fail to do so.

2.4.3 Prosody of givenness

Prosody plays an important role for tracking referents in spoken discourse. It goes beyond a dichotomous distinction between givenness and deaccentuation, on the one hand, and newness and pitch accentuation, on the other hand. As proposed by Prince (1981) and Chafe (1994), givenness can be understood as a continuum on which given referents are located at one extreme of the continuum while new referents are located at the other extreme. Accessible referents are located somewhere in between (Chafe 1994). Many studies on spoken discourse in German have shown that the degree to which a referent is given is reflected by the accent type with which it is realised. Thus, different accent types are considered to be information structurally meaningful by signalling different degrees of givenness. Accent types which exhibit higher prosodic prominence were often realised with less activated discourse referents. In contrast, accent types which exhibit lower prosodic prominence were often realised with more activated discourse referents (Baumann et al. 2015). Findings on the role of prosody in reference tracking not only show its importance but also point towards a more fine-grained distinction of the intermediate givenness category of accessibility (Baumann & Grice 2006). A selected number of studies on German about the role of prosodic prominence in spoken discourse will be discussed in the following.

Baumann et al. (2015) conducted a study on speech production in which the prosodic realisation of four different degrees of givenness were investigated. Native speakers of German were asked to read aloud different stories which contained referents that were either *new, accessible, given* or *given-displaced*.[11] Throughout the different stories, the carrier sentences as well as the target referents were held constant. All carrier sentences started with a personal pronoun of the grammatical subject role which was followed by a finite phrasal verb, which, in turn, was followed by the target referent and ended in a verbal particle which was part of the aforementioned phrasal verb. An example item is provided in (17) (Baumann et al. 2015: 15).

(17) Er schaut sich die Nina an.

The results showed differences in the type and placement of the nuclear pitch accent. The nuclear pitch accent was placed either on the target referent *die Nina* or on the sentence-final verb particle *an*. In the latter case, the target

[11] Baumann et al. (2015) used the *RefLex* notions of *r-unused, r-bridging, r-given* and *r-given-displaced* (Riester & Baumann 2017).

referent either received a prenuclear pitch accent or was deaccented altogether. Overall, new referents almost always received the highest prosodic prominence profile. They were preferably realised with L+H* or H* in nuclear position. Referents that were more given often received prenuclear accents or no accents at all. As a results, detecting a preferred pitch accent type for given and accessible referents was more difficult. However, accessible referents received significantly more often nuclear pitch accents than given referents. The differences in type and placement of the nuclear pitch accents could be ascribed to the differences in the information structural processing of the referents: More given referents were marked by lower prosodic prominence while less given ones were marked by higher prosodic prominence.

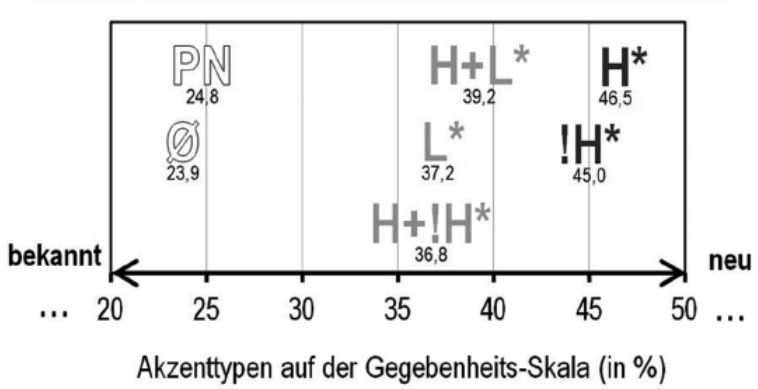

Figure 11: Prosodic marking of givenness in German (Baumann et al. 2015: 26).

Baumann et al. (2015) conducted two studies on speech perception in order to investigate whether listeners are able to clearly identify different degrees of givenness of referents solely by means of their prosodic marking. Both perception experiments used speech material that was elicited during the speech production study mentioned above. In one of the perception studies, participants were presented with the target sentence out of context. For this, the target sentence was extracted from the story. After listening to the sentence, they were asked to evaluate the degree of familiarity or *Bekanntheitsgrad* of the referent in the sentence. The results are shown in Figure 11 (Baumann et al. 2015: 26). Referents which were realised with a prenuclear pitch accent (PN) or no pitch accent at all (∅) were perceived to be significantly more given. In contrast, the presence of a nuclear pitch accent was strongly associated with lower degrees of givenness. Referents with higher (H*) or downstepped (!H*) pitch were interpreted to be least given. Lower pitch, either with (H+L* / H+!H*) or without (L*) an early peak, was interpreted to be neither fully given nor fully new. The difference between the perceived degrees of givenness, where no accent marks the highest degree of givenness, low pitch accents

mark an intermediate degree of givenness and high pitch accents mark the least amount of givenness, was statistically significant. These findings further suggest that not only the pitch height on the accented syllable but also the leading tone before the accented syllable plays a crucial role in the givenness interpretation of referents: A leading early peak (H+!H*) was found to add to the perceived degree of givenness while a lack thereof (!H*) led to a decrease.

The other perception experiment was conducted to test the appropriateness of the prosodic marking on the referent in the context of a specific information structural background. Thus, unlike in the previous perception experiment, the target sentences were embedded in a story based on which the target referents were either new, accessible, given or given-displaced. After listening to the whole story, participants were asked to evaluate the appropriateness of the intonation of the target sentence. The results showed that nuclear pitch accents on referents were perceived to be more appropriate the less activated the referents were in the discourse. In contrast, prenuclear pitch accents or the absence of accents on the referent were perceived to be more appropriate the more activated the referents were in the discourse. Importantly, the difference for given and accessible referents was found to be statistically significant. Apart from tendencies, no significant effects were found for new referents. This was attributed to the design of the items. Items which were designed for new referents started with a general question, which probably allowed for a wide variety of prosodic realisations as every element in the story was part of the broad focus domain. In contrast, the items for given and accessible referents mentioned the referents or the anchor of the referents early in the story because of which the target referents were part of the background domain allowing only for a limited number of prosodic realisations.

(18) a. Zu zweit sollen die Kinder im Biologie-Unterricht getrocknetes Obst analysieren... Sie werfen die Rosine weg.
b. Thomas darf heute im Zoo seinen Lieblingsaffen füttern... Er steckt sich die Banane ein.
c. Die Eltern sind sich unsicher mit einem neuen Medikament, das sie vom Arzt für ihr Kind bekommen haben... Sie rufen Dr. Bahber an.
d. Der Oberarzt und seine Kollegen brauchen für ihr Team einen neuen Orthopäden... Sie laden Dr. Bieber ein.

A closer look at the items for the accessible referents reveals that Baumann et al. (2015) did not control for the type of accessibility of the referent. For instance, *getrocknetes Obst* and *die Rosine* in (18a) clearly constitute a hyper-

nym–hyponym relation. In contrast, the relation between *sein Lieblingsaffe* and *die Banane* in (18b) is not that clear. Moreover, the weak definite in (18c) encodes a uniqueness relation between *vom Arzt* and *Dr. Bahber* (Schwarz 2009). In (18d), however, the indefinite article in *einen neuen Orthopäden* signals that *Dr. Bieber* is chosen from an unlimited group of orthopaedists (Gundel et al. 1993).

Baumann & Grice (2006) postulate that accessibility cannot be treated as a uniform category. They conducted a speech perception experiment in which they investigated the appropriateness of different pitch accent types (∅ / H+L* / H*) that are associated with the different degrees of givenness for eight different types of accessibility in German. The different types of accessibility included *textual displacement, scenarios,* symmetrical lexical relations such as *synonymy* and *converseness,* asymmetrical lexical relations such as *hyponymy–hypernymy*, *hypernymy–hyponymy* and *meronymy* including both *whole–part* as well as *part–whole* relations. Example noun pairs for each type are provided in (19).

(19)

a.	Textual displacement:	a house. . . the house
b.	Synonymy:	lift–elevator
c.	Converseness:	teacher–pupil
d.	Hypernymy–hyponymy:	fruit–orange
	Hyponymy–hypernymy:	dog–pet
e.	Whole–part:	book–page
	Part–whole:	roof–house
f.	Scenario:	bus–driver

Textual displacement differs from all other accessibility types in that the referent has been explicitly mentioned earlier in the discourse. Due to the referent not being rementioned for three clauses (Yule 1981), the referent recedes from a former active state to a semi-active one (Chafe 1994). The residual accessibility types belong to the category of inferential accessibility (Prince 1981). Instead of being explicitly mentioned in the discourse, they are semi-activated by an explicitly mentioned anchor referent. The relation between the anchor and the referent is either logical or lexical as exemplified from (19a) to (19e), or it is established by a culture-specific *scenario* (Sanford and Garrod 1981) as exemplified in (19f). Note that the scenario semi-activates not only one referent but a set of referents which are conceptually related to the scenario. An example would be the scenario of a court which semi-activates the concept of a judge, a lawyer, a juror etc. (Baumann & Grice 2006: 1645).

Table 1: Preferred accent types for different types of accessibility. '≫' indicates a highly significant preference, '>' indicates a significant preference and '=' indicates no significant preference (Baumann & Grice 2006: 1650).

Type of accessibility	Pitch accent type preferences
Converseness	No accent ≫ H+L* > H*
Part-whole	No accent ≫ H+L* ≫ H*
Synonymy	No accent ≫ H+L* > H*
Hyponym-hypernym	No accent ≫ H+L* ≫ H*
Hypernym-hyponym	No accent ≫ H+L* > H*
Textually displaced	H+L* = no accent ≫ H*
Whole-part	H+L* ≫ H* = no accent
Scenario	H+L* > H* = no accent

Baumann & Grice (2006) found a correlation between pitch accent type and accessibility type. Their results are shown in Table 1 (Baumann & Grice 2006: 1650). Deaccenting the accessible referent was strongly preferred if coreferential interpretations could be accommodated. This was the case for synonyms (*lift–elevator*), which established identity relations (van Deemter 1992), as well as part–whole (*roof–house*) and hyponym–hypernym (*dog–pet*) relations where the subordinate referent (*roof, dog*) established the *concept-givenness* of the superordinate one (*house, pet*) (van Deemter 1992). Interestingly, hyperonym–hyponym (*fruit–orange*) relations were also preferably deaccented. This could be attributed to the unique identifiability of the subordinate referent (*orange*) which was denoted by a definite noun phrase (Gundel et al. 1993). In case of converseness, deaccentuation was preferred because the accessible member (*pupil*) of the converseness relation could be interpreted to be coreferential with another referent in the story which was denoted by a proper name. Crucially, the second most preferred accent type was H+L*, an accent type that is preferred for accessible referents, while the least acceptable accent type was H*, an accent type which often marks new referents (Baumann et al. 2015). The significant difference in the prosodic marking suggests, albeit, indirectly, that H+L* has an intermediate status.

For textually displaced referents (*a house. . . the house*) no accent and H+L* were equally appropriate while H* was inappropriate. This suggests that after three clauses the degree of givenness of the referent (*the house*) has decreased slightly allowing for both realisations. Referents from whole-part (*book–page*) relations as well as scenarios (*bus–driver*) were preferably marked with H+L*. Marking these referents (*page / driver*) with H* or no accent at all was highly dispreferred. The results suggest that—provided that no givenness interpret-

ation can be accommodated—H+L* serves as an "accessibility accent" (Baumann & Grice 2006: 1654) and can be regarded as the most appropriate accent type for referents that are semi-activated from either whole–part relations or scenarios.

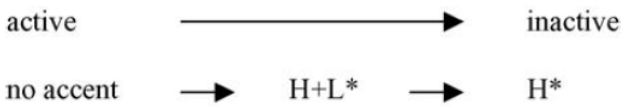

Figure 12: Prosodic marking of activation degrees in German (Baumann & Grice 2006: 1655).

A number of studies have shown that a referent's activation state can be understood to be gradient in nature and that different pitch accent types and positions indicate meaningful differences between different activation states. This is shown in Figure 12 where no accent signals an active state, H+L* an intermediate or semi-active state and H* an inactive state. A gradual downgrading of a referent's degree of activation (active > inactive) is reflected in the gradual upgrading of the referent's prosodic prominence (no accent < H+L* < H*). H+L* has an earlier peak than H* because of which H* is perceived to be more prosodically prominent and more appropriate for newer referents. H+L*, on the other hand, contributes to the impression of increased givenness.

2.5 Implicit causality

One research area in reference tracking is concerned with causal relations in discourse. Causal relations between consecutive discourse utterances (U_n, U_{n+1}) can be established by the verb contained in the first utterance (U_n). This is exemplified in (20). Verbs which establish causal relations are called *implicit causality verbs* (IC verbs). IC verbs are defined as "transitive verbs with two animate arguments characterized by the particular property of triggering explanations focusing systematically on one of the two arguments when followed by a because clause" (Bott & Solstad 2014: 214). In other words, the proportion of explanations attributed to either the first referent (REF_1) as exemplified in (20a) or the second referent (REF_2) as exemplified in (20b) expressed in U_2 varies as a function of the type of IC verb contained in U_1. The preference for either REF_1 or REF_2 which is brought about by the IC verb is called *implicit causality bias* (IC bias).

(20) a. $[[\text{Mary}]_{\text{REF}_1}$ fascinated $[\text{John}]_{\text{REF}_2}]_{U_1}$ [because <u>she</u> always knew what to say.$]_{U_2}$

b. $[[\text{Mary}]_{\text{REF}_1}$ admired $[\text{John}]_{\text{REF}_2}]_{\text{U}_1}$ [because he always knew what to say.$]_{\text{U}_2}$

The IC bias has direct implications for the ranking of referents for becoming the topic of the subsequent utterance. Many researchers have attributed the ranking of referents specifically to the lexical characteristics of the referents. Those characteristics include the referent's grammatical role (subject > object) (Grosz et al. 1995), their thematic role assigned by the binding verb (agent > patient) (Brown & Fish 1983; Garvey & Caramazza 1974; Malle 2002) and even their socio-cultural role assigned by their natural gender (male > female) (Ferstl et al. 2011). Despite the amount of research, findings have been contradicting each other immensely. According to Kehler et al. (2008), favouring heuristic over discourse structural factors combined with methodological errors such as "the degree of freedom afforded in the selection of stimuli" (Kehler et al. 2008: 38) fed into the body of contradicting findings.

A recently published paper by Bott & Solstad (2014) provides a novel account which attributes the IC bias not only to the lexical information available in the first utterance but also to the explanation type occurring in the immediately following utterance. Kehler et al. (2008) argued that a referent's argument structure alone is insufficient in predicting which referent is to occur in the subsequent utterance. They were the first to observe systematic dependency relations between coreference and explanation type. Based on their findings, Bott & Solstad (2014) postulated a semantic-propositional theory which provided an explanation for the dependency relation between IC verbs and the explanations relating to one of the binding referents. By conducting experiments for German and Norwegian, Bott & Solstad (2014) provide empirical evidence for the universal applicability of their theory.

(21) a. Mary_{S} fascinated John_{E} because she danced very well.
b. Mary_{E} admired John_{S} because he sang beautifully.

In their semantic-propositional theory, Bott & Solstad (2014) take into account different types of IC verbs including, among others, *stimulus-experiencer verbs* (SE verbs) and *experiencer-stimulus verbs* (ES verbs) (Gernsbacher & Hargreaves 1988). As the order within the names suggests, SE verbs assign stimulus roles to the subject argument and experiencer roles to the object argument while ES verbs assign experiencer roles to the subject argument and stimulus roles to the object argument.[12] Both IC verb types belong to the class of *psychological verbs* and denote states (Brown & Fish 1983). For both psy-

[12]This is the case for most West-Germanic languages like English and German.

chological verbs, it is the stimulus argument and not the experiencer argument which strongly attracts the IC bias (Brown & Fish 1983). Thus, utterances which contain SE verbs tend to be continued with the subject argument while utterances which contain ES verbs tend to be continued with the object argument. Examples for utterances containing SE verbs and ES verbs as well as their respective IC biases are given in (21).

(22) a. John_{S} annoyed Mary_{E}.
b. John_{S} annoyed Mary_{E} because he sang loudly.

Bott & Solstad (2014) attribute the stimulus argument's property of attracting the IC bias to the stimulus argument's underlying semantic representation. On the surface, both stimulus and experiencer arguments are represented in terms of noun phrases. Their underlying representations, however, differ in that the stimulus constitutes a proposition while the experiencer constitutes an entity. Lexical entries of SE verbs like *annoy* entail the mental state denoted by the IC verb, a *causer* which is propositional in nature and a *causee* which is merely a sentient being. Assigning noun phrases to the causer and the causee will specify the causee sufficiently but not the causer. Being propositional in nature, the causer carries an additional empty slot for events. Leaving this slot unspecified will evoke a sense that specific information is missing. This is exemplified by the utterance in (22). (22a) lacks information about the source of *Mary's* annoyance which is rooted in *John*. The IC bias, or better yet, the *stimulus bias*, follows from a "general processing preference for not leaving missing content unspecified" (Bott & Solstad 2014: 219). As can be seen in (22b), modifying the matrix clause with a *because* clause will naturally result in specifying "which underspecified property of John it was that caused annoyance in Mary" (Bott & Solstad 2014: 222).

(23) SE VERB:
a. John annoyed Mary (because he sang loudly).
b. It annoyed Mary that John sang loudly.

(24) ES VERB:
a. Mary hated John (because he sang loudly).
b. Mary hated it that John sang loudly.

One piece of evidence in favour of the propositional analysis of the stimulus argument is that stimulus arguments alternate with *that* clauses which are regarded to be "indisputably propositional in nature" (Bott & Solstad 2014: 223). This is the case for both SE verbs and ES verbs as can be seen in (23)

and (24). Another piece of evidence is that causal relations hold even without the explicit mention of the *because* conjunction (Kehler et al. 2008). Just like (25a) (Bott & Solstad 2014: 213), which contains a *because* connector, (25b) (Bott & Solstad 2014: 213), which lacks a *because* connector, yields a causal relation between the first and the second utterance.

(25) a. Mary_{S} fascinated John_{E} because she always knew what to say.
b. Mary_{S} fascinated John_{E}. She always knew what to say.

The different types of IC verbs strongly correlate with different types of explanations (Bott & Solstad 2014). For instance, SE verbs and ES verbs trigger simple causes but not externally or internally anchored reasons. The different explanation types are exemplified in (26) (Bott & Solstad 2014: 220). Simple causes are causes of events, states or attitudinal states which "never involve volition or agentivity on the side of the causing individual" (Bott & Solstad 2014: 220). Thus, the disturbance caused by *John* in (26a) is unintentional and rather results as an epiphenomenon of *John's* action. In contrast, internally and externally anchored reasons are causes of attitudinal states which involve intentions on the part of the attitude-bearer (Solstad 2010). While causes of externally anchored reasons are rooted outside the attitude-bearer's mind, causes of internally anchored reasons are rooted within the attitude-bearer. In (26b-i), the disturbance caused by *John* is rooted in an external event in which *Mary* damages *John's* bike. In (26b-ii), on the other hand, the disturbance caused by *John* is rooted in his internal state of anger.

(26) a. SIMPLE CAUSE:
John disturbed Mary because he was making lots of noise.

b. (i) EXTERNALLY ANCHORED REASON:
John disturbed Mary because she had damaged his bike.
(ii) INTERNALLY ANCHORED REASON:
John disturbed Mary because he was very angry at her.

2.6 Summary

Many different factors contribute to the way referents are structured in spoken discourse. Although using different terms, both Grosz et al. (1995) and Chafe (1994; 1996) claim that the amount of cognitive effort placed upon the interlocutors plays a crucial role in the processing of discourse referents. Referents that cause less inference load (Grosz et al. 1995) or activation cost (Chafe

1994; 1996) are easier to process and tend to be taken up again in the ongoing discourse more frequently. The amount of cognitive processing is highly dependent on the referent's degree of givenness. Given referents are already part of the common ground of the speaker and the listener because of which they tend to require the least amount of activation cost. In contrast, new referents need to be activated in the consciousness of the listener, which requires the highest amount of activation cost. A referent's degree of givenness is signalled by the linguistic expression used to identify it. Researchers are, however, divided with regard to the amount of contribution different linguistic factors pay towards the referent in becoming the next backward-looking center or topic of the subsequent utterance. On the on hand, Grosz et al. (1995) claim that continuing the discourse with the (utterance-initial) subject argument requires the least amount of cognitive effort while shifting the topic by continuing the discourse with the (utterance-non-initial) object argument is more strenuous to the listener's mind. Bott & Solstad (2014), on the other hand, attribute the likelihood of a referent to be taken up again to the thematic role that the referents were assigned by the implicit causality verb. In case of psychological verbs, referents become the topic of the subsequent utterance if they are of the stimulus role. Another factor that contributes to processing differences is the lexical realisation of the discourse referent. According to Gundel et al. (1993), Givón (1983) and Ariel (1988; 1990; 1991; 1996), the type of the lexical expression determines the degree of discourse prominence of the referent. The higher the discourse prominence of a referent the higher its probability of being continued in the discourse. According to Gundel et al. (1993), definite constructions are more discourse prominent than their indefinite counterparts. Givón (1983) and Ariel (1988; 1990; 1991; 1996) further arrange different types of definite descriptions into a hierarchy of discourse prominence. Less lexico-phonological coding material is associated with higher discourse prominence which is supposed to facilitate the cognitive accessibility of the referent. This is in line with Baumann & Grice (2006) who claim that in German more given referents bear no pitch accent while less given ones tend to be realised with fully-fledged pitch accents. Thus, referents are considered to be highly given if they are maximally discourse prominent and minimally prosodic prominent. This would be the case for a referent denoted by a deaccented pronoun. In contrast, referents are considered to be least given if they are minimally discourse prominent and maximally prosodic prominent, which would hold for a referent denoted by an indefinite noun phrase realised with an L+H*. Not only does L+H* constitute the most prominent accent type of German, it also functions as a contrastive focus marker (Mücke & Grice 2004) which makes alternatives

to the referent more prominent, thereby diminishing the probability of being taken up in the subsequent utterance even further (Braun 2015).

3 Story continuation task

3.1 Research question and hypotheses

The story continuation task was designed to test the effect of syntactic, semantic, prosodic and information structural factors on reference tracking in German. All stories were controlled for with respect to the referent's syntactic role as well as their thematic role which was assigned by the binding implicit causality verb (Bott & Solstad 2014). All stories contained a subject argument in utterance-initial topic position and an object argument in utterance-medial non-topic position (Halliday 1967; 1970). In half of the stories implicit causality verbs assigned a stimulus role to the subject and an experiencer role to the object while in the other half of the stories, an experiencer role was assigned to the subject and a stimulus role to the object. This mirror-inverted assignment of thematic roles onto referents with fixed grammatical roles allowed for the examination of two opposing hypotheses. On the one hand, Grosz et al. (1995) and Kaiser (2011) claim that referents of grammatical subject roles rank higher than referents of grammatical object roles. Thus, an overall preference for subjects, i.e. a *subject bias*, is to be expected irrespective of the referent's thematic role. Bott & Solstad (2014) and Kehler et al. (2008), on the other hand, deny the prevalence of a heuristic subject bias. Instead, they emphasise the importance of thematic roles in discourse. According to Bott & Solstad (2014), there is an overall preference for specifying stimulus arguments, i.e. a *stimulus bias*, but not experiencer arguments with simple causes. In other words, stimulus arguments rank higher than experiencer arguments, irrespective of their grammatical role.

In addition to grammatical and thematic role, all stories were controlled for their prosodic realisation. The prosodic realisation of the referent in the non-topic position varied with respect to three different pitch accent types: A falling pitch accent (H+L*), a rising pitch accent (L+H*) and the absence of a pitch accent, deaccentuation (∅).[13] All stories were set up as scenarios (Sanford & Garrod 1981) which semi-activated the referent in non-topic position (Chafe 1994). Thus, H+L* was considered to be the default realisation while ∅ and L+H* constituted mismatches between the activation state and

[13]For the sake of simplicity, ∅ will be regarded as an accent type along with H+L* and L+H*.

the prosodic realisation of the referent. The chosen pitch accent differed in terms of their prosodic prominence profiles in that ∅ was the least prominent, L+H* the most prominent while H+L* was neither, because of which H+L* was considered to have an intermediate prosodic prominence (Baumann & Röhr 2015).

As ∅ and L+H* have different functions, they were assumed to affect the prevalent bias in different ways. ∅ and L+H* were assumed to increase the cognitive effort required to process the stories (Grosz et al. 1995). As a result, participants would increasingly rely on the existing biases in order to reduce the cognitive burden placed on them. That is, the less appropriate the prosodic accent the stronger the subject bias or stimulus bias, respectively. For the stories used in this experiment, L+H* was expected to be more appropriate than ∅ because another information structural interpretation could be accommodated. That is, L+H* was assumed to activate explicit alternatives to the focused non-topic referent (Braun 2015; Mücke & Grice 2014). As the referent in topic position was the only explicit alternative available, L+H* was assumed to make the referent in topic position more prominent.

All hypotheses will be described in terms of a referent's *topic shift potential* (Baumann & von Heusinger 2016). The topic shift potential describes the potential of a referent in utterance-medial, non-topic position to be shifted to the utterance-initial, topic position in the following utterance. In this thesis, however, topic shift potential will be described as the potential of the *bias attracting referent* in becoming the topic of the subsequent utterance which includes not only topic shifts but also topic continuations. The referent's topic shift potential will be regarded in relative terms which will be symbolised by the '>' sign, such that everything to its left indicates a higher potential while everything to its right indicates a lower potential of becoming the next topic. Equal potentials will be symbolised by the '=' sign.

H1: Subject bias. If the referent is the subject of an utterance U_n, it is most likely to become the topic of the subsequent utterance U_{n+1}. If the referent is the object of an utterance U_n, it is unlikely to become the topic of the subsequent utterance U_{n+1}. Overall, no topic shifts are to be expected.

SUBJECT > OBJECT
Lisa$_{\text{SUBJ}}$ hat den Trainer$_{\text{OBJ}}$ beeindruckt. *Sie...*
Maja$_{\text{SUBJ}}$ hat den Sänger$_{\text{OBJ}}$ vergöttert. *Sie...*

H2: Stimulus bias. If, in an utterance U_n, the referent in subject position is assigned the stimulus role and the referent in object position is assigned the experiencer role, the referent in subject position is most likely to become the topic of the subsequent utterance U_{n+1}; relatively more topic continuations are to be expected. If, in an utterance U_n, the referent in subject position is assigned the experiencer role and the referent in object position is assigned the stimulus role, the referent in object position is most likely to become the topic of the subsequent utterance U_{n+1}; relatively more topic shifts are to be expected. In both cases, the subsequent utterance U_{n+1} will contain a simple cause which further elaborates on the stimulus argument.

SE = ES
Lisa$_{S}$ hat den Trainer$_{E}$ beeindruckt. *Sie...*
Maja$_{E}$ hat den Sänger$_{S}$ vergöttert. *Er...*

H3a: Subject bias & accent type. Differences in the prosodic realisation of the referent in non-topic object position will not overwrite the heuristic subject bias. However, different accent types will affect the referent's topic shift potential to different degrees. The subject bias will be strongest for deaccented referents as processing difficulties evoked by the inappropriateness of ∅ will be compensated for by relying on the subject bias even more. Thus, deaccented object referents will undergo the least amount of topic shifts. The subject bias will be less pronounced for referents marked by H+L*. Thus, referents marked by H+L* will undergo more topic shifts than deaccented referents. Referents marked by L+H* will activate the object referent—the only explicit alternative to the subject referent. Thus, referents marked by L+H* will undergo the highest amount of topic shifts without overwriting the overall subject bias.

∅ > H+L* > L+H*
Lisa$_{SUBJ}$ hat den Trainer$_{OBJ}$(∅) beeindruckt. *Sie...*
Maja$_{SUBJ}$ hat den Sänger$_{OBJ}$(∅) vergöttert. *Sie...*
Lisa$_{SUBJ}$ hat den Trainer$_{OBJ}$(H+L*) beeindruckt. *Sie...*
Maja$_{SUBJ}$ hat den Sänger$_{OBJ}$(H+L*) vergöttert. *Sie...*
Lisa$_{SUBJ}$ hat den Trainer$_{OBJ}$(L+H*) beeindruckt. *Sie...*
Maja$_{SUBJ}$ hat den Sänger$_{OBJ}$(L+H*) vergöttert. *Sie...*

H3b: Stimulus bias & accent type. The stimulus bias for SE verbs will exhibit a similar pattern as the subject bias. Deaccenting the experiencer will strengthen the stimulus bias resulting in the least amount of topic shifts. For referents marked by H+L*, more topic shifts will be observable than for referents marked by ∅. L+H* will activate the alternative experiencer referent which will reduce the cognitive effort to take up the experiencer. Thus, L+H* will result in the highest amount of topic shifts but it will not have the power to overwrite the stimulus bias.

SE(∅) > SE(H+L*) > SE(L+H*)
Lisa$_S$ hat den Trainer$_E$(∅) beeindruckt. *Sie...*
Lisa$_S$ hat den Trainer$_E$(H+L*) beeindruckt. *Sie...*
Lisa$_S$ hat den Trainer$_E$(L+H*) beeindruckt. *Sie...*

ES verbs will show a different bias pattern than SE verbs. The default H+L* realisation of the stimulus will require the least amount of cognitive effort. Thus, H+L* will lead to the highest amount of topic shifts. Marking the stimulus with L+H* will facilitate the cognitive accessibility to the experiencer—the only explicit alternative to the stimulus. As a result, less topic shifts and more topic continuations will be expected. Deaccenting the stimulus will require the highest amount of cognitive effort. The inappropriateness of ∅ will create a conflict between the stimulus bias which favours the stimulus over the experiencer and the mismatch which disfavours the stimulus as the topic of the subsequent utterance. As a consequence, no difference in the amount of topic shifts and topic continuations will be expected.

ES(H+L*) > ES(L+H*) > ES(∅)
Maja$_E$ hat den Sänger$_S$(H+L*) vergöttert. *Er...*
Maja$_E$ hat den Sänger$_S$(L+H*) vergöttert. *Er...*
Maja$_E$ hat den Sänger$_S$(∅) vergöttert. *Er... / Sie...*

3.2 Method

3.2.1 Material

The story continuation task employed a 2×3 factorial design with two levels of IC verb types (SE vs. ES) and three levels of accent types (H+L* vs. L+H* vs. ∅). This resulted in sixty test items which were divided across three lists. Recordings took place in the I*f*L-Phonetics Laboratory at the University of Cologne. They were done at 44kHz/16bit using a DAT-recorder and a condenser microphone. All test items were recorded inside a soundproof booth and read at a normal rate by a trained female phonetician.

(27) SE ITEM
Freitagnachtmittag war das Fitnessstudio mal wieder total überfüllt.
Lisa hat den Trainer beeindruckt.

(28) ES ITEM
Die Oper am Samstag war komplett ausverkauft.
Maja hat den Sänger vergöttert.

Each test item was a story consisting of a sequence of two utterances as exemplified in (27) for SE items and in (28) for ES items. The first utterance introduced a scenario with an open setting. The second utterance consisted of two arguments which were bound by a transitive IC verb. The first argument was in subject position and information structurally new. It referred to a female individual which was denoted by a proper name. The second argument was in direct object position and accessible via the scenario in the first utterance. It referred to a male individual which was denoted by a definite noun phrase. Both arguments were disyllabic and had a strong-weak metrical stress pattern.

The IC verb was in perfect tense. It consisted of the conjugated auxiliary verb *haben*, which was placed between the two arguments, and the past participle form of the IC verb, which was placed at the end of the utterance. One advantage of using perfect tense as opposed to past tense was having a past participle at the end of the utterance which prevented prosodic clustering on the second argument. Another advantage was that it constitutes the narrative form in German. Hence, it was expected to facilitate the elicitation of story continuations.

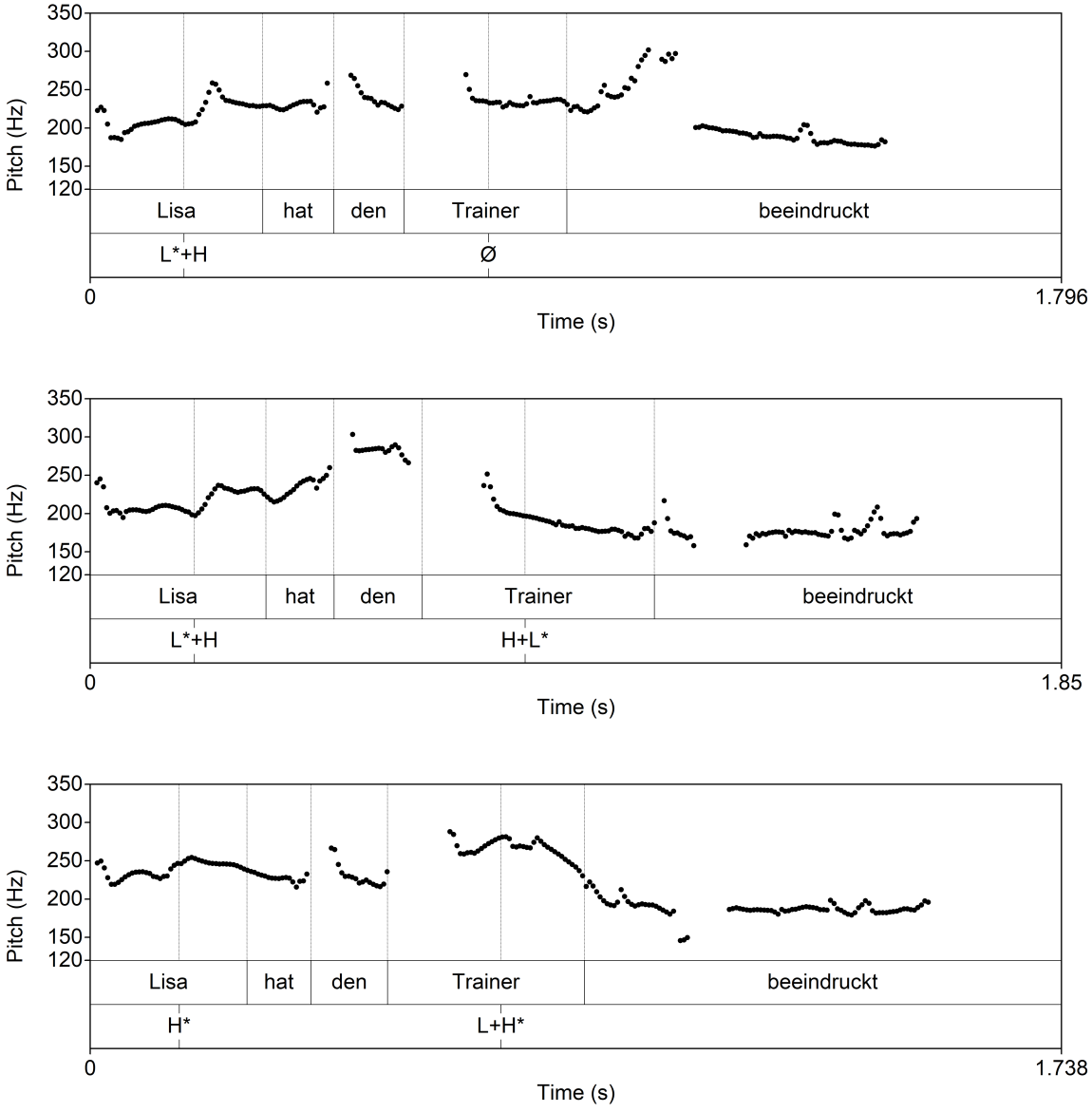

Figure 13: F0 contours of stories containing the SE verb *beeindrucken*. The non-topic argument *den Trainer* is either realised with ∅ (top), H+L* (middle) or L+H* (bottom).

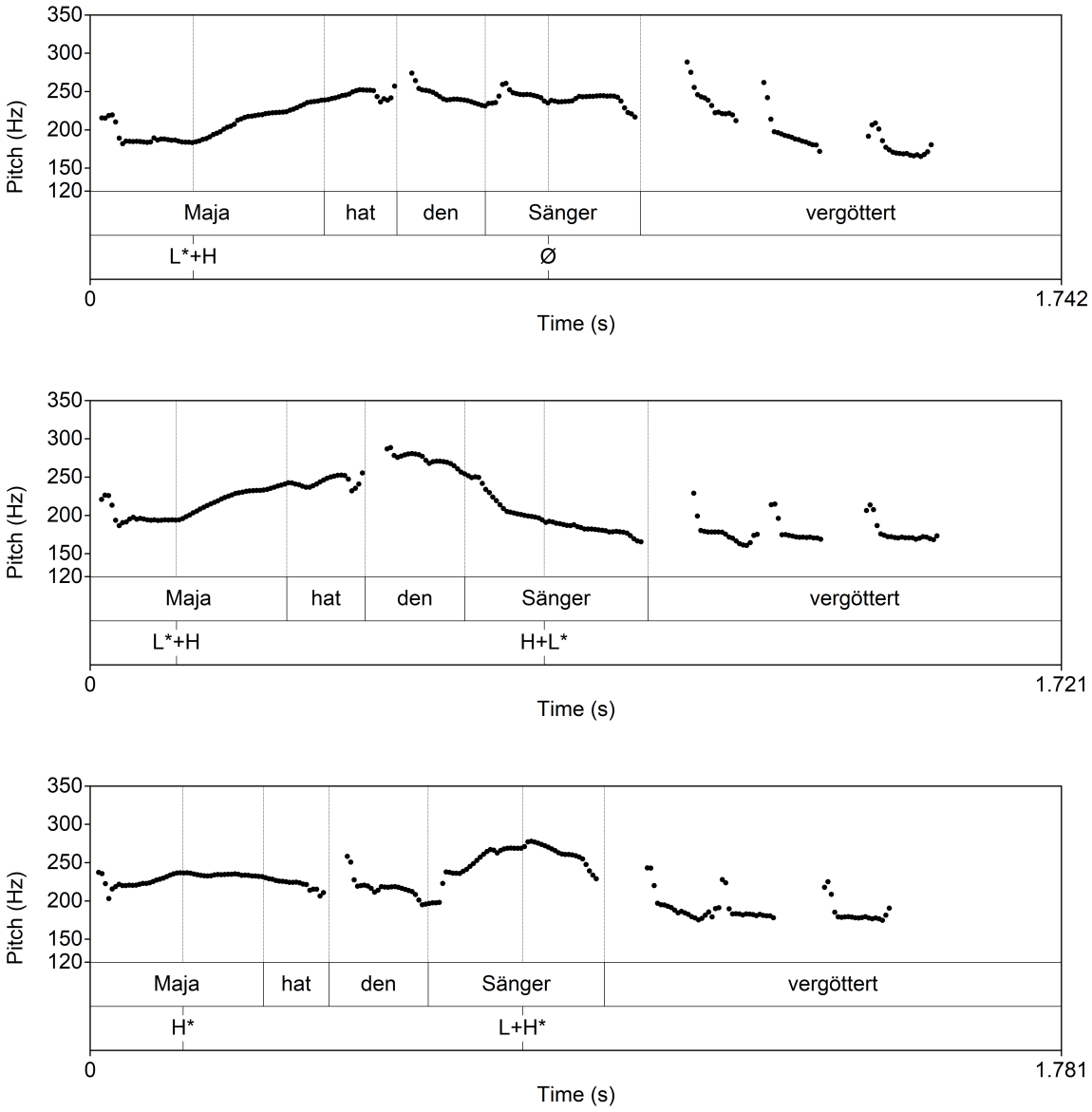

Figure 14: F0 contours of stories containing the ES verb *vergöttern*. The non-topic argument *den Sänger* is either realised with ∅ (top), H+L* (middle) or L+H* (bottom).

3.2.2 Procedure

The story continuation task was conducted online using SoSci Survey (Leiner 2019). It was made available via the link http://www.soscisurvey.de. Participants were able to do the experiment from any web-enabled device of their choice. After opening the link, a new browser window with requirements and instructions for the story continuation task appeared. Requirements included a good internet connection, a quiet environment and earphones or loud speakers. Participants were instructed to listen to twenty short stories. Their task was to continue each story with at least three sentences and at most five sentences. In order to prevent long and elaborate story continuations, participants were told to keep their continuations short. Moreover, they were informed that the quality of their story continuations was not evaluated. The task was self-paced and did not include a training phase prior to the test phase. During the test phase, participants were able to play back the items as often as required.

3.2.3 Participants

Nineteen native speakers of German participated in the story continuation task without receiving any monetary compensation. Ten participants were female and nine participants were male. Their mean age was 25.8 years. All participants had *Abitur* and nine of them had an additional university degree. Neither hearing nor major visual impairments were reported.

3.2.4 Annotation

All story continuations were annotated manually using Microsoft Excel (Microsoft Office 365 ProPlus, Version 16.0.12430.20172). Annotations were limited to referential expressions that occurred in the first utterance. For annotation purposes, an utterance was defined as a clause which contained maximally one conjugated verb. The first utterance was either a matrix clause as exemplified by the underlined sentence in (29) or a subordinate clause as exemplified in (30).

(29) MATRIX CLAUSE

a. [Sie]$_{\text{REF}_1}$ konnte mehr Gewichte heben als sonst.

b. [Der Trainer]$_{\text{REF}_2}$ freute sich darüber, dass Lisa seit der Vertragsunterzeichnung so konsequent das Fitnessstudio besucht hatte.

(30) SUBORDINATE CLAUSE

Obwohl [sie]$_{\text{REF}_1}$ erst seit wenigen Wochen Judo macht, hat sie eine hervorragende Technik gezeigt.

In line with Halliday (1967; 1970), potential re-mentions of the first (REF_1) and the second referent (REF_2) were annotated with regard to their order of mention within the utterance. Thus, a referent's order of mention was used as an indicator to determine whether referents underwent topic continuations, topic retentions or topic shift (Grosz et al. 1995).

(31) TOPIC CONTINUATION

a. Mara hat den Jäger gefürchtet.
[Mara]$_{\text{REF}_1}$ hat den Schuss in naher Umgebung einschlagen hören.
REF_1=SUBJECT, REF_2=NONE
ORDER OF MENTION=REF_1

b. Lisa hat den Trainer beeindruckt.
Mit nur wenigen Hilfsmitteln gelang es [ihr]$_{\text{REF}_1}$ ...
REF_1=INDIRECT OBJECT
ORDER OF MENTION=REF_1

c. Frieda hat den Onkel verachtet.
[Frieda]$_{\text{REF}_1}$ hat [ihren Onkel]$_{\text{REF}_2}$ beobachtet, ...
REF_1=SUBJECT, REF_2=DIRECT OBJECT
ORDER OF MENTION=REF_1, REF_2

Topic continuation was defined as the first mention of REF_1 in the subsequent utterance. As can be seen in (31), this was independent of the grammatical role of REF_1. The presence or absence of another referent was also not taken into consideration.

(32) TOPIC RETENTION
Lea hat den Rentner verabscheut.
[Der Rentner]$_{\text{REF}_2}$ hat genauso auch [Lea]$_{\text{REF}_1}$ verabscheut.
REF_1=DIRECT OBJECT, REF_2=SUBJECT
ORDER OF MENTION=REF_2, REF_1

Topic retention was defined as the first mention of REF_2 and the second mention of REF_1 in the subsequent utterance. This is illustrated in (32).

(33) TOPIC SHIFT

a. Vera hat den Gärtner entdeckt.
[Er]$_{\text{REF}_2}$ war am Blumengießen.
REF_1=NONE, REF_2=SUBJECT
ORDER OF MENTION=REF_2

b. Jana hat den Maler inspiriert.
Mit ihrem auffälligen Kleidungsstil. . . kam [ihm]$_{\text{REF}_2}$ die Idee.
REF_1=NONE, REF_2=INDIRECT OBJECT
ORDER OF MENTION=REF_2

c. Frieda hat den Onkel verachtet.
[Der Onkel]$_{\text{REF}_2}$ beleidigte [ihre Schwester]$_{\text{REF}_3}$.
REF_1=NONE, REF_2=SUBJECT, REF_3
ORDER OF MENTION=REF_2, REF_3

Topic shift was defined as the first mention of REF_2. As can be seen in (33), this was independent of the grammatical role of the second referent. The presence or absence of another referent was not taken into consideration.

(34) NO TOPIC
Pia hat den Bäcker gemocht.
Die Bäckerei roch nach frischen Brötchen.
REF_1=NONE, REF_2=NONE
ORDER OF MENTION=NONE

No topic was defined as the absence of a referent in the newly produced utterance. This is exemplified in (34).

3.3 Data analysis

Story continuations from one participant needed to be excluded from the statistical analysis because he misunderstood the task. Story continuations which contained no topics were also omitted. Thus, all statistical analyses were based on (380 – 20 'excluded participant' – 56 'no topic' =) 304 story continuations produced by eighteen participants. Further inspections of the data revealed that 46.7% of the data were topic continuations while 41.5% were topic shifts. Topic retentions amounted to 11.8% because of which they were regrouped as topic shifts resulting in a total of 53.3% of topic shifts.

The hypotheses, introduced in §3.1, were tested using two different statistical analyses. H1 and H2 were tested performing Chi-square tests of independence using the `chisq.test()` function in R (R Core Team 2019). Chi-square tests were performed to test whether the topic shift potential was dependent of the subject bias or the stimulus bias, respectively. H3a and H3b were tested performing Cochran-Mantel-Haenszel (CMH) tests using the `mantelhaen.test()` function from the `DescTools` (Signorell 2020) package in R (R Core Team 2019). CMH tests were conducted to test the effect of a third variable, namely accent type (∅ vs. H+L* vs. L+H*), on the relation between the subject bias and topic shift potential, on the one hand, and the stimulus bias and topic shift potential, on the other hand.

3.4 Results

Table 2: Percentages of topic continuations and topic shifts with regard to the subject and stimulus bias.

	TOPIC CONT	TOPIC SHIFT
SUBJ	46.7%	53.3%
SE	65.7%	34.3%
ES	29.3%	70.7%

The Chi-square test revealed that the relation between the subject bias and topic shift potential was not significant ($\chi^2(1)$=1.32, p=0.25). In contrast, the relation between stimulus bias and topic shift potential was found to be significant ($\chi^2(1)$=33.9, p=5.8×10^9). These findings reject the existence of a heuristic subject bias as hypothesised by H1 but provide support for the existence of a stimulus bias as hypothesised by H2. Table 2 shows the percentages of topic continuations and topic shifts with regard to the subject and the stimulus bias. Overall, less topic continuations (46.7%) than topic shifts (53.3%) were produced. For SE verbs, topic continuations occurred significantly more often than topic shifts (65.7% > 34.5%). For ES, there were significantly more topic shifts than topic continuations (70.7% > 29.3%).

As H1 was proven to be false, H3a could be rejected as well. The remaining hypothesis, H3b, was tested by performing a CMS test. Accent type was found to be a significant factor in the relation between the stimulus bias and topic shift potential ($\chi^2(1)$=32.79, p=1×10^8). As can be seen for SE verbs in Table 3, decreasing the prosodic prominence of the experiencer led to more topic continuations (72% > 66.7% > 54.2%) while increasing the prosodic prominence led to more topic shifts (45.8% > 33.3% > 28%). For ES verbs, it was the

Table 3: Percentages of topic continuations and topic shifts with regard to the stimulus bias for different accent types.

		TOPIC CONT	TOPIC SHIFT
SE	∅	72.0%	28.0%
	H+L*	66.7%	33.3%
	L+H*	54.2%	45.8%
ES	∅	39.1%	60.9%
	H+L*	21.4%	78.6%
	L+H*	32.1%	67.9%

default realisation with H+L* which led to the highest amount of topic shifts (78.6%) and the lowest amount of topic continuations (21.4%). More topic shifts were observed with contrastively marked stimulus referents than with deaccented ones (67.9% > 60.9%). Likewise, deaccentuation resulted in more topic continuations than contrastive focus (39.1% > 32.1%). As illustrated by the ranking below, these results confirm H3b concerning the influence of cognitive effort mediated by different accent types on SE as well as for ES verbs.

Predicted pattern: SE(∅) > SE(H+L*) > SE(L+H*)
Results: SE(∅) > SE(H+L*) > SE(L+H*)

Predicted pattern: ES(H+L*) > ES(L+H*) > ES(∅)
Results: ES(H+L*) > ES(L+H*) > ES(∅)

4 Discussion

4.1 Ranking of referents

The influence of the different linguistic factors on the ranking of discourse referents was assessed indirectly by examining which of the referents were taken up again in the subsequent discourse utterance. While grammatical role failed to predict the topic-to-be, implicit causality and prosodic prominence were identified as good topic predictors. This is illustrated in Figure 15.

The results indicate that discourse referents are constantly competing one another for the first rank. REF_1 was given a competitive head start by being supplied with the highly discourse prominent utterance-initial subject position. REF_2 was given the less discourse prominent utterance-medial object position. This asymmetry in discourse prominence was amplified by the choice of linguistic expressions. While REF_1 was realised with a proper name (*Lisa* / *Maja*),

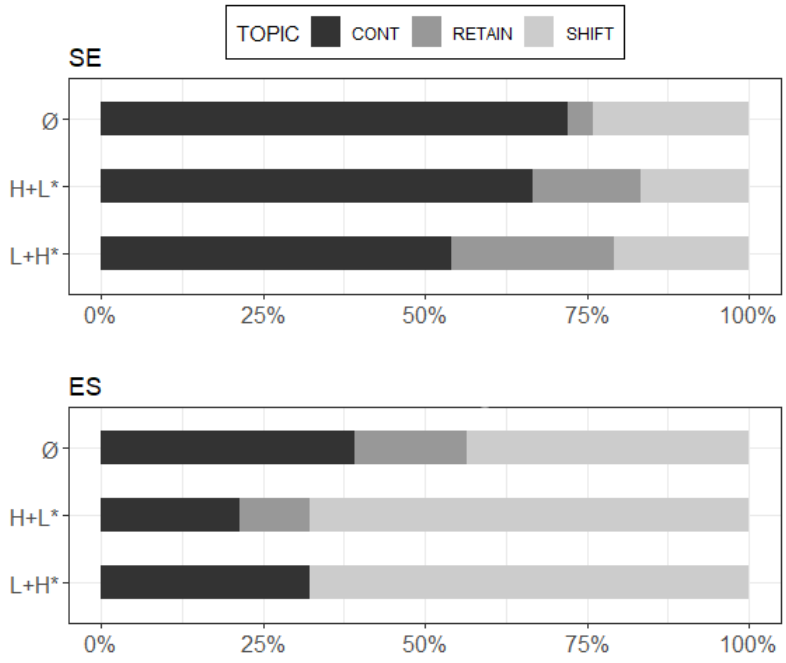

Figure 15: Ratio between topic continuation, topic retention and topic shift as a function of accent type for SE (top) and ES items (bottom).

REF_2 was realised with a definite noun phrase (*den Trainer* / *den Sänger*). According to the Accessibility Theory by Ariel (1991), *first names* rank higher than *short definite descriptions* and are, therefore, cognitively more accessible to the listener.

When comparing the activation states of both referents, REF_2 was more activated than REF_1. As a result, REF_2 was accessible from the setting, hence requiring less activation cost. In contrast, REF_1 is discourse-new and needs to be activated from an inactive state. This might explain why the percentage of topic continuations for SE verbs was only 66.7% while the percentage of topic shifts (including topic retentions) for ES verbs amounted to 78.6% in the default condition.

Overall, the contrastive focus marker, L+H*, affected the stimulus bias such that explicit alternatives to the bias attracting referent were made more prominent. For both SE and ES verbs, the explicit alternative was always the experiencer. Thus, focus marking activated the experiencer which is against the stimulus bias. As a result, more topic shifts were observed for SE verbs (45.8%) while more topic continuations were observed for ES verbs (32.1%). In both cases, however, contrastive focus marking was not strong enough to overwrite the stimulus bias.

Deaccentuation primarily functions as a givenness marker. Deaccented referents are considered to be highly discourse prominent while at the same time being prosodically non-prominent. In the stories used in the story continuation task, a givenness interpretation of REF_2 was inhibited due to the setting of the story. As a result, participants were confronted with a prosodic match for REF_1 and a prosodic mismatch for REF_2. The results show that participants frequently opted for REF_1 which was more appropriate in terms of both prosodic realisation and activation state. The mismatch between deaccentuation, on the one hand, and semi-activation, on the other hand, of REF_2 promoted a bias towards REF_1 in addition to the existing stimulus bias. As a consequence, the topic continuation rate for both SE and ES was highest: SE verbs exhibited 72% and ES verbs exhibited 39.1% of topic continuations.

4.2 Choice of linguistic expression

The choice of linguistic expression indicates to what degree a referent was activated in the ongoing discourse. §2.3 introduced three discourse prominence hierarchies by Ariel (1991), Givón (1983) and Gundel et al. (1993). Apart from some differences, they all claim that, among the definite descriptions, "unstressed" pronouns rank highest in the hierarchy for which they are maximally discourse prominent. In addition, linguistic expressions rank higher if they are able to identify unequivocally the referent in question (Ariel 1991). Thus, proper names are considered to be more discourse prominent than definite noun phrases (Ariel 1991). The question arises whether different degrees of prosodic prominence constrained the availability of certain linguistic expressions such that lower degrees of prosodic prominence correlate with more discourse prominent linguistic expressions while higher degrees of prosodic prominence correlate with less discourse prominent linguistic expressions.

Figure 16 shows the percentages of the linguistic expressions that were chosen for all REF_2 (*den Trainer / den Sänger*) which underwent topic retentions or topic shifts. In the subsequent utterance, REF_2 was continued to varying extents with either a personal pronoun, a definite noun phrase or a demonstrative pronoun. Irrespective of accent type, REF_2 was preferably taken up with a personal pronoun. For deaccented referents, however, the preference for personal pronouns was strongest. Although to a lesser extent, REF_2 that was marked by H+L* was preferably taken up by a definite noun phrase while those marked by L+H* tended to be taken up by demonstrative pronouns.

The distribution of linguistic form over accent type does not show a reverse correlation between prosodic prominence marked by different accent types and

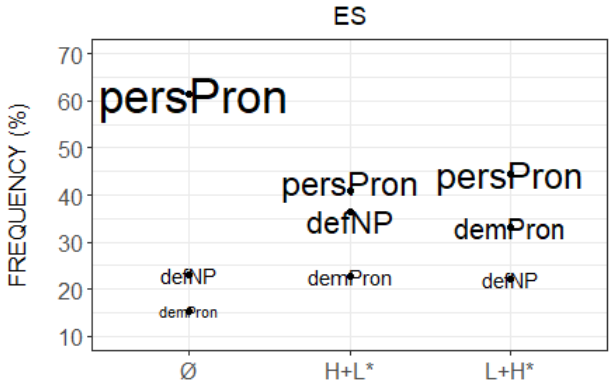

Figure 16: Frequency of linguistic expressions over accent type for topic shifts and topic retentions for ES items.

discourse prominence marked by choice of linguistic expression. H+L*, which has a lower prosodic prominence profile than L+H*, licensed the use of definite noun phrases which are less discourse prominent. Referents which were realised with L+H*—the most prominent accent type—tended to be taken up by demonstrative pronouns which rank higher than definite noun phrases. Although no statistical analysis was conducted, these results indicate that theories on discourse prominence need to refine their hierarchies by taking different accent types into consideration and going beyond a mere distinction between "stressed" and "unstressed" phonological coding.

4.3 IC bias and explanation type

For both SE and ES verbs, the default realisation with H+L* exhibited a relatively low stimulus bias. While for ES verbs, topic shifts amounted to 78.6%, SE verbs elicited only 66.7% of topic continuations. This could be attributed to the type of explanation given by the participants. The stories used in the experiment were very short because of which information on the referents were quite scarce. As a result, participants often adopted a strategy of beginning their story continuations by providing some background information about the experiencer. Consequently, the stimulus bias was masked even more.

(35) Die Oper am Samstag war komplett ausverkauft. [Maja]$_{\text{E}}$ hat [den Sänger]$_{\text{S}}$ vergöttert. [Sie hatte gerade noch so ein Ticket ergattern können...]$_{\text{BACKGROUND}}$

Further analyses revealed that simple causes accounted for the majority of all story continuations which contained at least one referent (65.5%). However,

background causes (Bott & Solstad 2014) as the one exemplified in (35) made up 34.5%. According to Bott & Solstad (2014), background causes are provided in order to facilitate the situation described by the IC verb. They differ from simple causes in that they do not specify the reason why the experiencer was put in a certain state.

(36) a. SIMPLE CAUSE:
John frightened Mary because he suddenly screamed.

b. BACKGROUND CAUSE:
Felix frightened Vanessa because she didn't hear him coming.

The difference between simple and background causes is exemplified in (36) (Bott & Solstad 2014: 9). While the *because* clause in (36a) specifies the screaming of John as the cause of Mary's state of being frightened, the *because* clause in (36b) merely specifies the background for Mary being frightened. The background cause does not specify what John did that led to Mary's unfortunate state.

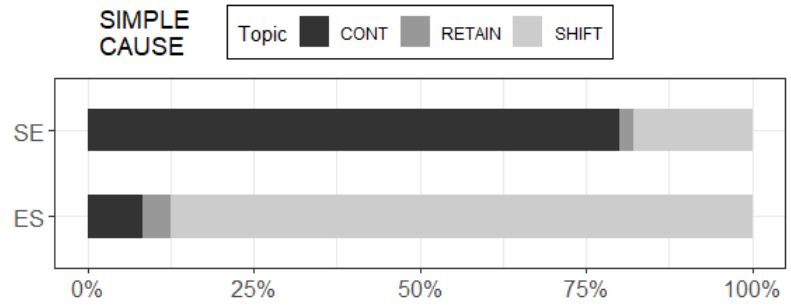

Figure 17: Ratio between topic continuations, topic retentions and topic shifts as a function IC verb type for simple causes.

Figure 17 shows the relative proportions of topic continuations, retentions and shifts for simple causes only. The preference for the stimulus argument was over 75% for both SE and ES items. In comparison, a joint analysis of simple and background causes revealed a weaker stimulus bias (see Table 2). A separate analysis of the background causes revealed a bias towards the experiencer argument. This is illustrated in Figure 18. While SE verbs often led to a high proportion of topic shifts and topic retentions, ES verbs were found to trigger topic continuations.

The stories used in this study triggered a relatively high amount of background causes. This was unexpected as Bott & Solstad (2014), who used the same SE and ES verbs, found an insignificantly small amount of background causes. The overall high proportion, however, indicates that participants had difficulties in performing the task. Participants were subjected to very short stories which ended in a *full-stop* prompt. Adding an intermediate sentence that provides some background information on the experiencer might have prevented the high proportion of background causes. Moreover, using a *because* prompt instead of a *full-stop* prompt might have signalled to the participants that the stories were underspecified with regard to their causal content better. An improved version of a story is exemplified in (37).

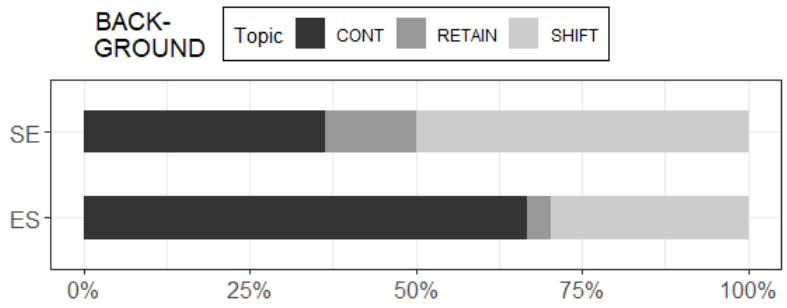

Figure 18: Ratio between topic continuations, topic retentions and topic shifts as a function of IC verb type for background causes.

(37) Die Oper am Samstag war komplett ausverkauft. (SETTING)
Maja hatte gerade noch so ein Ticket ergattern können. (BACKGROUND)
Sie hat den Sänger vergöttert, weil... (TARGET)

5 Conclusion

This thesis investigated how different linguistic factors from various linguistic fields, including syntax, semantics, prosody and information structure, affected the ranking of referents in becoming the topic of the subsequent utterance in a highly controlled story continuation task in German. Native speakers of German were asked to listen to and continue stories in which the non-topic referent was accessible via the setting the story took place. Referents were either bound by SE verbs, which assigned the stimulus role to the topical subject argument, or by ES verbs, which assigned the stimulus role to the non-topical object argument. Some of the stories were prosodically preferred

as the non-topical referent was realised with an H+L*. Others were designed to induce processing difficulties, either by placing a contrastive pitch accent on the target referent or by deaccenting it completely. Mismatches in prosody forced participants to accommodate other information structural interpretations or increasingly rely on the prevalent bias.

The results of the story continuation favoured a stimulus bias over a heuristic subject bias. In line with Bott & Solstad (2014) and Kehler et al. (2008), this thesis provides further evidence against a heuristic subject bias as proposed by Grosz et al. (1995). Instead, it shows the importance of lexico-semantic factors like implicit causality for information structuring in discourse.

Different prosodic realisations either strengthened or weakened the stimulus bias without ever overwriting it. The cognitive effort placed on the listener, which was mediated by a match or mismatch between information structure and accent type, was found to be a reliable predictor of the subsequent topic of the ongoing discourse. Less cognitive effort was found to increase the potential of the bias attracting referent in becoming the topic of the subsequent utterance. In contrast, more cognitive effort led to a gradual decrease of the bias. Interestingly, this was irrespective of the relative prosodic prominence of the different accent types as accent types were interpreted with regard to the information structural context.

This thesis showed that lexico-semantic as well as prosody play an integral part in information structure and information processing. For this reason, both factors must neither be turned a blind eye to nor a deaf ear to any longer! The findings of this thesis strongly suggest that future studies on discourse planning should control for the prosodic and lexico-semantic realisations of their stimuli.

Appendix

Table 4: SE items.

Setting	Target
Freitagnachtmittag war das Fitnessstudio mal wieder total überfüllt.	Lisa hat den Trainer beeindruckt.
Die Stimmung während des gestrigen Schulunterrichts war äußerst angespannt.	Ina hat den Lehrer enttäuscht.
Freitagabend ging es im besten Restaurant der Stadt hektisch zu.	Laura hat den Kellner verwirrt.
Kurz nach Feierabend war der Bus extrem voll.	Eva hat den Fahrer irritiert.
Durch die großen Fenster des Ateliers kam heute wieder sehr viel Sonnenlicht herein.	Jana hat den Maler inspiriert.
Nach Klärung vieler Differenzen war die Scheidung fast vollzogen.	Lara hat den Anwalt überrascht.
Beim diesjährigen Schachturnier ging es hochkonzentriert zu.	Lola hat den Gegner verblüfft.
Im Seniorenheim waren Monotonie und Einsamkeit an der Tagesordnung.	Nora hat den Pfleger gelangweilt.
Während der Verhandlung wurde ein schwieriger Fall diskutiert.	Nina hat den Richter fasziniert.
Der morgendliche Judo-Kurs war sehr gut besucht.	Klara hat den Meister entzückt.

Table 5: ES items.

Setting	Target
Beim jährlichen Familienfest wurde wieder viel erzählt.	Frieda hat den Onkel verachtet.
Die Oper am Samstag war komplett ausverkauft.	Maja hat den Sänger vergöttert.
Gestern Abend fand eine Podiumsdiskussion zur Gender-Debatte statt.	Lena hat den Redner respektiert.
Montagmorgen lief das Fußball-Training bereits auf Hochtouren.	Paula hat den Stürmer bewundert.
Während der Wanderung durch den Wald waren Schüsse zu hören.	Mara hat den Jäger gefürchtet.
Bei der alljährlichen Kaffeefahrt gab es wieder viel Gedränge.	Lea hat den Rentner verabscheut.
Am Samstag wurde in Köln ein Tango-Schnupperkurs angeboten.	Mia hat den Tänzer beneidet.
Der Gottesdienst am Sonntag war besonders festlich.	Sara hat den Pfarrer vermisst.
Aus der Konditorei an der Ecke roch es mal wieder herrlich.	Pia hat den Bäcker gemocht.
Der nahegelegene Stadtpark war wieder besonders gut gepflegt.	Vera hat den Gärtner entdeckt.

References

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

Ariel, Mira. 1990. *Accessing noun phrase antecedents.* London: Routledge.

Ariel, Mira. 1991. The function of accessibility in a theory of grammar. *Journal of Pragmatics* 16. 443–463.

Ariel, Mira. 1996. Referring expressions and the +/- coreference distinction. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility.* Amsterdam: John Benjamins Pusblishing Company. 13–36.

Baumann, Stefan. 2006. *The intonation of givenness: Evidence from German. Linguistische Arbeiten 508.* Tübingen: Max Niemeyer Verlag.

Baumann, Stefan. 2008. Degrees of givenness and their prosodic marking. In Riehl, Claudia Maria & Astrid Rothe (eds.), *Was ist linguistische Evidenz? Kolloquium des Zentrums Sprachenvielfalt und Mehrsprachigkeit, November 2006.* Aachen: Shaker Verlag. 35–55.

Baumann, Stefan & Martine Grice. 2006. The intonation of accessibility. *Journal of Pragmatics* 38(10). 1636–1657.

Baumann, Stefan & Arndt Riester. 2013. Coreference, lexical givenness and prosody in German. *Lingua* 136. 16–37.

Baumann, Stefan & Christine T. Röhr. 2015. The perceptual prominence of pitch accent types in German. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences.* Glasgow: The University of Glasgow. Paper number 298. 1–5.

Baumann, Stefan, Christine T. Röhr & Martine Grice. 2015. Prosodische (De-)Kodierung des Informationsstatus im Deutschen. *Zeitschrift für Sprachwissenschaft* 34(1). Berlin: De Gruyter. 1–42.

Baumann, Stefan & Klaus von Heusinger. 2016. *The role of accent types in referent tracking.* Talk presented at Prosody and Information Structure (PINS), Stuttgart March 22, 2016.

Beckman, Mary E. & Julia Hirschberg. 1994. *The ToBI annotation conventions: Manuscript and accompanying speech material.* Ohio State University.

Beckman, Mary E., Julia Hirschberg & Stefanie Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing.* Oxford: Oxford University Press. 9–54.

Boersma, Paul & David Weenink. 2019. Praat: Doing phonetics by computer (Version 6.1.08) [Computer program]. http://www.praat.org/.

Bolinger, Dwight L. 1961. Contrastive accent and contrastive stress. *Language* 37(1). 83–96.

Bott, Oliver & Torgrim Solstad. 2014. From verbs to discourse: A novel account of implicit causality. In Barbara Hemforth, Barbara Mertins & Cathrine Fabricius-Hansen (eds.), *Conversational implicatures in anaphora resolution: Alternative constructions and referring expressions.* Cham: Springer. 213–251.

Braun, Bettina. 2006. Phonetics and phonology of thematic contrast in German. *Language and Speech* 49(4). 451–493.

Braun, Bettina. 2015. What causes the activation of contrastive alternatives, the size of focus domain or pitch accent type? In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings 18th International Congress of Phonetic Sciences.* Glasgow: The University of Glasgow. Paper 379. 1–5.

Brown, Roger & Deborah Fish. 1983. The psychological causality implicit in language. *Cognition* 14(3). 237–273.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li (ed.), *Subject and topic.* New York, NY: Academic Press. 25–56.

Chafe, Wallace L. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing.* Chicago: University of Chicago Press.

Chafe, Wallace L. 1996. Inferring identifiability and accessibility. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility.* Amsterdam: John Benjamins Pusblishing Company. 37–46.

Chomsky, Noam. 1969. *Deep structure, surface structure and semantic interpretation.* Bloomington: Indiana University Linguistics Club.

Clark, Herbert H. 1977. Bridging. In Philip N. Johnson-Laird & Peter C. Wason (eds.), *Thinking: Readings in cognitive science.* Cambridge: Cambridge University Press. 411–420.

Clark, Herbert H. 1996. *Using language.* Cambridge: Cambridge University Press.

Dudenredaktion. n.d. In Duden online. Retrieved from https://www.duden.de/rechtschreibung/Katar.

Ferstl, Evelyn C., Alan Garnham, & Christina Manouilidou. 2011. Implicit causality bias in English: A corpus of 300 verbs. *Behavior Research Methods* 43(1). 124–135.

Garvey, Catherine, & Alfonso Caramazza. 1974. Implicit causality in verbs. *Linguistic Inquiry* 5(3). 459–464.

Gernsbacher, Morton A. & David J. Hargreaves. 1988. Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language* 27. 699–717.

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

Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics Vol. 3: Speech acts.* New York, NY: Academic Press. 41-58.

Grice, Martine & Stefan Baumann. 2002. Deutsche Intonation und GToBI. *Linguistische Berichte* 191. 267–298.

Grice, Martine & Stefan Baumann. 2016. Intonation in der Lautsprache: Tonale Analyse. In Ulrike Domahs & Beatrice Primus (eds.), *Handbuch Laut, Gebärde, Buchstabe.* Berlin: De Gruyter. 84-105.

Grice, Martine, Mathias Reyelt, Ralf Benzmüller, Jörg Mayer & Anton Batliner 1996. Consistency in transcription and labelling of German intonation with GToBI. In H. Timothy Bunnell & William Idsardi (eds.), *Proceedings of the 4th International Conference on Spoken Language Processing.* Philadelphia, PA: University of Delaware. 1716–1719.

Grice, Martine, Stefan Baumann & Ralf Benzmüller. 2005. German intonation in autosegmental-metrical phonology. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press. 55–83.

Groenendijk, Jeroen A. G. 1999. The logic of interrogation: Classical version. In Tanya Matthews & Devon Strolovitch (eds.), *Proceedings of the 9th Semantics and Linguistic Theory Conference*. Santa Cruz, CA: University of California. 109–126.

Grosz, Barbara. J., Aravind K. Joshi & Scott Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21(2). 202–225.

Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274-307.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation.* Cambridge: Cambridge University Press.

Halliday, Michael A. K. 1967. Notes on transitivity and theme in English. Part 2. *Journal of Linguistics* 3(2). 199–244.

Halliday, Michael A. K. 1970. Language structure and language function. In John Lyons (ed.), *New horizons in linguistics*. Harmondsworth: Penguin Books. 140–165.

Heim, Irene R. 1982. *The semantics of definite and indefinite noun phrases.* Doctoral dissertation. Amherst, MA: University of Massachusetts.

Hockett, Charles. 1958. *A course in modern linguistics.* New York, NY: McMillan Company.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar.* Cambridge, MA: Massachusetts Institute of Technology Press.

Jadgfeld, Nils & Stefan Baumann. 2011. Order effects on the perception of relative prominence. In Wai-Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences.* Hong Kong: City University of Hong Kong. 958–961.

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

Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen A. G. Groenendijk, Theo M. V. Janssen & Martin B. J. Stokhof (eds.), *Formal methods in the study of language.* Amsterdam: Mathematisch Centrum. 277–322.

Karttunen, Lauri. 1974. Discourse referents. In James D. McCawley (ed.), *Syntax and semantics 7: Notes from the linguistic underground.* New York, NY: Academic Press. 363–385.

Katz, Jonah & Elisabeth O. Selkirk. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language* 87(4). 771–816.

Kehler, Andrew, Laura Kertz, Hannah Rohde & Jeffrey L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics* 25(1). 1–44.

Krifka, Manfred & Renate Musan. 2012. Information structure: Overview and linguistic issues. In Manfred Krifka & Renate Musan (eds.), *The expression of information structure.* Berlin: De Gruyter Mouton. 1–44.

Ladd, Robert. 2008. *Intonational phonology.* Cambridge: Cambridge University Press.

Lambrecht, Knud. 1994. *Information structure and sentence form. Topic, focus and the mental representation of discourse referents.* Cambridge: Cambridge University Press.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8(1). 339–359.

Leiner, Dominik. J. 2019. SoSci Survey (Version 3.2.03-i) [Computer software]. http://www.soscisurvey.com.

Malle, Bertram. 2002. Verbs of interpersonal causality and the folk theory of mind and behavior. In Masayoshi Shibatan (ed.), *The grammar of causation and interpersonal manipulation.* Amsterdam: John Benjamins. 57–83.

Merin, Arthur. 1994. Algebra of elementary social acts. In Savas L. Tsohatzidis (ed.), *Foundations of speech act theory: Philosophical and linguistic perspectives.* London: Routledge. 234–266.

Molnár, Valéria & Hélène Vinckel-Roisin. 2019. Discourse topic vs. sentence topic. Exploiting the right periphery of German verb-second sentences. In

Valéria Molnár, Verner Egerland & Susanne Winkler (eds.), *Architecture of topics*. Berlin, Boston, MA: De Gruyter. 293–334.

Mücke, Doris & Martine Grice. 2014. The effect of focus marking on supralaryngeal articulation—is it mediated by accentuation? *Journal of Phonetics* 44. 47–61.

Prince, Ellen F. 1981. Towards a taxonomy of given-new information. In Peter Cole (ed.), *Radical Pragmatics.* New York, NY: Academic Press. 223–256.

R Core Team. 2019. R: A language and environment for statistical computing (Version 1.2.5001) [Computer program]. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophika* 27(1). 53–94.

Reyelt, Matthias, Martine Grice, Ralf Benzmüller, Jörg Mayer & Anton Batliner. 1996. Prosodische Etikettierung des Deutschen mit ToBI. In Dafydd Gibbon (ed.), *Natural language and speech technology, results of the third KONVENS conference.* Berlin: Mouton de Gruyter. 144–155.

Riester, Arndt & Stefan Baumann. 2017. *The RefLex scheme–annotation guidelines.* SinSpeC 14: Working Papers of the SFB 732 "Incremental Specification in Context". Stuttgart: Online Publikationsverbund der Universität Stuttgart.

Rochemont, Michael. 2016. Givenness. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*. Oxford: Oxford University Press. 41–63.

Rochemont, Michael. 2019. Topics and givenness. In Valéria Molnár, Verner Egerland & Susanne Winkler (eds.), *Architecture of topics*. Berlin: De Gruyter. 47–66.

Röhr, Christine T. & Stefan Baumann. 2010. Prosodic marking of information status in German. *Proceedings of the 5th International Conference on Speech Prosody*. Chicago, IL. Paper 100019. 1–4.

Sanford, Anthony J. & Simon C. Garrod. 1981. *Understanding written language: Explorations of comprehension beyond the sentence.* Chichester, NY: John Wiley.

Schumacher, Petra B., Jana Backhaus & Manuel Dangl. 2015. Backward- and forward-looking potential of anaphors. *Frontiers in Psychology* 6. Article 1746. 1–14.

Signorell, Andri. 2020. *DescTools: Tools for descriptive statistics.* R package version 0.99.32 https://CRAN.R-project.org/package=DescTools.

Selkirk, Elisabeth O. 1995. Sentence prosody: Intonation, stress and phrasing. In John A. Goldsmith (ed.), *Handbook of phonological theory.* Cambridge, MA: Blackwell. 550–569.

Solstad, Torgrim. 2010. Some new observations on 'because (of)'. In Maria Aloni, Harald Bastiaanse, Tikitu de Jager & Katrin Schulz (eds.), *Logic, Language and Meaning: 17th Amsterdam Colloquium.* Berlin: Springer. 436–445.

Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton K. Munitz & Peter K. Unger (eds.), *Semantics and Philosophy.* New York, NY: New York University Press. 197–214.

Uhmann, Susanne. 1991. *Fokusphonologie: Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie.* Tübingen: Max Niemeyer Verlag.

van Deemter, Kees. 1992. Towards a generalization of anaphora. *Journal of Semantics* 9(1). 27-51.

Watson, Duane G., Michael K. Tanenhaus & Christine A. Gunlogson. 2008. Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science* 32. 1232–1244.

Wickham, Hadley. 2009. *ggplot2: Elegant graphics for data analysis.* New York, NY: Springer Verlag.

Yule, George. 1981. New, current and displaced entity reference. *Lingua* 55. 41–52.