Prosodic cue weighting in sentence comprehension: Processing German case ambiguous structures

Dissertation

zur Erlangung des akademischen Grades
Doctor philosophiae (Dr. phil.)
an der Humanwissenschaftlichen Fakultät
der Universität Potsdam

eingereicht von
Anja Gollrad

Gutachter:
Prof. Dr. Caroline Féry
PD Dr. Frank Kügler

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-81954
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-81954

Contents

Acknowledgements

I would like to thank Caroline Féry and Frank Kügler for giving me the opportunity to write my dissertation on a topic that captured my interest already during my linguistic studies. Both agreed to be my advisors on the spur-of-the-moment and their support and guidance during this project was indispensable. The thesis resulted from my work as a Ph.D. student within the project "Prosody in Parsing" of the German Research Foundation (DFG) Priority program SPP 1234 "Sprachlautliche Kompetenz: zwischen Grammatik, Signalverarbeitung und neuronaler Aktivität", principle investigators Caroline Féry, Frank Kügler, and Shravan Vasishth.

I have been very fortunate to be part of the PiP project, and thank Caroline Féry, Frank Kügler, and Shravan Vasishth for a productive and supportive working atmosphere. Many thanks go to my colleague Gerrit Kentner with whom I shared an inspirational working environment in our joint office at the University of Potsdam in Golm from June 2006-September 2009, and also now over the Potsdam-Frankfurt/Main distance he continued to be a dear colleague and friend. I wish to thank Gerrit for his innumerable discussions on many aspects related to this work and beyond. Furthermore, I thank Umesh Patil and Stella Gryllia for the pleasant time that we spend together in Haus 35, where we were sharing offices at times.

Lyn Frazier's visit in the fall of 2008 in Potsdam established a corner stone in the process of developing and writing this thesis. Also, a discussion with Bod Ladd during the prosody conference ETAP 1 at Cornell University, where first parts of this work were presented, was very helpful and promotive as well. Merle Horne gave valuable comments during the Potsdam-Lund meeting in 2008, and Carlos Gussenhoven provided helpful contributions during his stay at the Workshop "Experimental Studies on Intonation", held in Potsdam in 2009. Further parts of this research were presented at the Speech Prosody 2010 in Chicago, where Amy Schafer invested a considerably amount of time in discussing frequency distributions and structural preferences.

A thanks goes out to all collegues and friends which whom I enjoyed the discussions on the experiments and their final interpretations. Thanks to Oliver Niehbuhr, Amy Schafer, Amanda Seidl, Claudia Männel, Petra Augurzky, Anja Arnhold, Susanne Genzel, Thomas Weskott, Anja-Kleemann-Krämer, Maria van Rell, Esther Sommerfeld, Fabian Schubö and Johannes Gollrad.

For many thoughtful, constructive and prompt comments on draft versions of this thesis I am very grateful to Caroline Féry, Frank Kügler, Gerrit Kentner, Natalie Boll-Avetisyan and Fabian Schubö. I greatly appreciated the efforts of Gerrit Kentner who spoke the stimuli for the perception experiments presented in Chapter 4, and Frank Kügler for his precise recordings for the perception experiments presented in Chapter 5. Frank, I also like to thank you for giving me the opportunity towards the end of the project to focus to intently on finishing this dissertation. Additionally, I would sincerely like to thank Gerrit Kentner, Fabian Schubö and Ruben van den Vijver who provided a substantial amount of time in discussing and contemplating on the details of modeling the perception data in an Optimality-Theoretic Framework.

For practical aspects related to the work, I like to thank Bernadett Smolibocki, Anne Schumacher and Franziska Reckling for their kind and helpful data collection and annotation of the production material. Stefan Schinkel helped me setting up the Latex layout and Julia Ritz provided considerably support for the TIGER-corpus request. Antje Sauermann and Sebastian Weinrich helped in questions related to the statistical analyses. I thank you all.

1

Preliminaries

1.1 Introduction

A universal characteristic of speech is that utterances are generally broken down phonologically into smaller phrases. These phrases can be further divided into smaller-sized constituents. Constituents of varying size, or «prosodic units» are typically marked by suprasegmental features such as intonational events and/or final lengthening. There is overall agreement that the prosodic marking of such constituents helps (a) listeners in identifying the underlying syntactic structure in general, and (b) resolving syntactic ambiguities such as *It's Johns turn to pay Emma*, where the sentence is usually understood to mean that John will pay *Emma*, when it is spoken as a single phrase with a prosodic boundary at the end of the word *Emma*. Is the same string of words spoken as two phrases, with prosodic boundaries at the end of the words *pay* and *Emma*, then the verb is interpreted to be intransitive, and the sentence is understood to be addressed to *Emma*.

Many studies have demonstrated that listeners can indeed rely on these kinds of prosodic differences to determine speakers intended meaning for such ambiguous utterances (e.g. Lehiste, 1973; Lehiste, Olive & Streeter, 1976; Price, Ostendorf, Shattuck-Hufnagel & Fong, 1991). However, there is less agreement about the *exact* prosodic information that is used by listeners in disambiguating the possible syntactic options of a syntactic ambiguity.

A way to address the problem is to examine the processing of utterances with temporary ambiguities by manipulating the prosodic correlates, such that the influence of each prosodic cue on sentence disambiguation can be attested. The results help us to decide how closely suprasegmental prosodic features are correlated with listeners' perception of prosodic phrase boundaries and further help us to derive a prosodic cue weighting on φ- and ι-boundaries in German. On the basis of existing models in phonological perception in Optimality Theory (Flemming, 1995; Boersma, 2009; Boersma & Hamann, 2009; Féry et al., 2009), a model of the perception of boundary related phonetic cues is proposed in this work. The results are relevant for a theory of perception of auditory inputs as they argue and account for the importance of different boundary cues related to different boundary sizes, and furthermore provide specific evidence for how one relevant phonetic cue guides the mapping of a phonetic input onto a phonological structure.

1.2 Structure of this dissertation

Chapter 2 introduces the theoretical background and shapens the research questions examined throughout this thesis. The assumptions about the universal characteristics of speech, in particular the partition of utterances into smaller phrases and their individual prosodic demarcation are made fully explicit. It is assumed that phrases are recursive and that tonal features in combination with durational cues guide the formation of prosodic phrases. In the second section, effects of prosodic phrasing on language comprehension are presented, and it is summarized, based on numerous experimental findings, that speakers use, and listeners are provided with, various prosodic cues to determine the underlying syntactic structure. Furthermore, the concept of prosodic cue weighting is introduced and discussed, and we see that languages like English, Dutch or Hindi strongly differ with respect to the prosodic cues used for disambiguation. Based on these results the conclusion is formulated that the role of prosody is rather language specific instead of universal and argues for the need to examine the various prosodic cues, provided by the speaker, that potentially guide the listener to syntactic features of the speech input in German. Chapter 2 concludes by outlining the aims of this dissertation.

Chapter 3 provides an overview of the experimental design and presents general aspects of the sentence material. In view of the hypothesis that German case ambiguous structures should be suitable for eliciting a cue weighting associated with prosodic phrase boundaries, the syntactic analysis of the ambiguity under scrutiny is introduced, thereby two syntactic structures are distinguished, a dative structure and a genitive structure. We discuss the syntactic complexity of these two structures, and show that the genitive structure is syntactically more complex than the alternative dative structure. We report a rating study on the effects of syntactic complexity on auditory language processing with a clear result of a preference for the dative structure over the genitive structure. In order to obtain supportive evidence for a preference for the dative structure which is attributable to the syntactic complexity rather than to frequency, we performed a corpus based frequency analysis of the two syntactic structures, rendering them compatible in frequency of occurrence. In this way, initial differences in the dynamics of processing between the two structures are relatable to syntactic rather than frequency distinctions.

Chapter 4 presents the first production experiment aiming to investigate the prosodic correlates and the corresponding prosodic phrasing of both sentence structures. In a series of six perception experiments, the impact of each individual prosodic correlate on sentence processing is separately examined and showed to influence parsing in such a way that a prosodic cue weighting on φ-boundaries can be established, with *prefinal lengthening* being the most important prosodic correlate.

Chapter 5 presents the second production experiment measuring the prosodic correlates of two specific readings of a syntactically ambiguous sentence containing an i-phrase boundary. A series of 5 perception experiments evaluated each prosodic cue on listeners' ability to disambiguate the sentences. Thereby the *quality of the boundary tone* (high versus low) preceding the ι-boundary appeared to be the most important cue for listeners to assign the correct syntactic analysis.

Chapter 6 takes up on the results that prosodic correlates show an individual influence of sentence disambiguation in dependence on the size of the boundary, and accounts for the observed pattern of cue weighting on φ-phrase and ι-phrase boundaries by making recourse to a grammatical Optimality-Theoretic model using boundary related phonetic constraints to model the mapping of a phonetic input representation onto a phonological structure.

Chapter 7 summarizes the main findings of the dissertation and suggests possible research endeavors with respect to the issues raised in this thesis.

2

Theoretical Background

2.1 Prosodic properties of speech

Spoken language is characterized not only by syntactic and semantic information, but also by a certain intonation and rhythmic structure, the prosody. Prosody is commonly defined as the set of suprasegmental features to convey postlexical or sentence-level pragmatic meanings (Ladd, 1996). The suprasegmental features are phonetically measurable in the acoustic speech signal and generally comprise (1) the fundamental frequency (f0) determined by the number of cycles the vocal cord open and close per second (measured in Hertz (Hz)), (2) the intensity or amplitude of a sequence which refers to the amount of energy, i.e subglottal pressure that is used during speech production (measured in decibel (db)), and (3) the duration which refers to the temporal aspect of speech, such as the duration of segments or pauses (measured in milliseconds (ms)) (Cruttenden (1986); Neppert & Pétursson (1986); Clark & Yallop (1995); Ladd (1996)).

These phonetic entities correspond to units in the perceptual domain. The perceptual counterpart of fundamental frequency is the tonal height and is referred to as pitch. Intensity/amplitude correlates with hearers' perception of loudness and the durational parameters are perceived as segmental, constituent or pause length. Table 2.1 shows the interplay of the acoustic parameters and their perceptual correlates. The systematic variation of prosody may signal discourse structure (Bolinger, 1972; Pierrehumbert & Hirschberg, 1990; Schafer, 1997; Ladd, 1996), sentence modes (declarative, imperative or interrogative) and paralinguistic information such as gender, age and the emotional state of the speaker.

Articulation	Acoustic	Perception
(1) vocal cord vibration	fundamental frequency (Hz)	pitch
(2) subglottal pressure	intensity /amplitude (db)	loudness
(3) timing of articulation	duration (ms)	length

Table 2.1: Prosodic parameters and their perceptive correlates

Prosodic correlates encode different pieces of linguistic information in different languages. German, English, and Dutch for example, have been traditionally classified as stress languages (Fry, 1955; Möbius, 1993). In these languages, prosody conveys lexical stress on the word level (Friedrich, 2003). On the sentence level, prosody signals, among other things, how prosodic constituents are grouped to each other. Such grouping information is referred to as phrasing. In psycholinguistic research there has been a growing interest in how phrasing correlates with syntactic structure, especially in the case of ambiguous sentences. The main focus of the present thesis will be on the the impact of prosodic constraints on syntactic ambiguity resolution and its effects on prosodic phrasing. The following section thus summarizes widely acknowledged phonological assumptions about the organization of prosodic constituents, focusing on the relationship between syntax and prosody.

Prosodic information is commonly referred to as suprasegmental information as it is considered to be located 'on top of' prosodic units such as syllables, words, or entire utterances. These units or domains are referred to as suprasegmental. In the framework of the suprasegmental phonology a strictly layered prosodic hierarchy for the linguistically relevant aspects of prosody is proposed (Selkirk, 1984; Nespor & Vogel, 1986). Traditionally, the prosodic hierarchy is referred to as Strict Layer Hypothesis (SLH), organizing prosodic constituents according to their respective size: smaller prosodic constituents are embedded in larger ones, such that syllables are embedded in words, prosodic phrases or utterances (1).

(1) The prosodic hierarchy (Nespor & Vogel, 1986) and (Selkirk, 1984)

Utterance (υ)
|
Intonational Phrase (ι)
|
Phonological Phrase (φ)
|
Prosodic Word (ω)
|
Food (Ft)
|
Syllabe (σ)

Models that assume such a hierarchical organization of prosodic constituents argue for an exhaustive parsing of each prosodic layer by the layer directly above it. Selkirk (1990) formulates this hypothesis as follows (P stands for any prosodic category):

(2) $P_n - P_{n-1}$

The representation of each prosodic category is hierarchically arranged, i.e. a higher or larger prosodic category Pn only dominates constituents of the next level down, Pn-1. For example, this means that in a scenario such as that described below in (3), the left boundary of the most leftward located phonological phrase (φ-phrase) coincides with the left boundary of the next upper Intonational Phrase (ι-phrase). The same holds true for the right boundary of the most rightward located φ-phrase which coincides with the right boundary of the next upper ι-phrase, respectively. This principle is referred to as exhaustiveness, claiming that no level of the prosodic hierarchy is supposed to be skipped, i.e. an ι-phrase does not immediately dominate a prosodic word (ω).

(3)

( )ι
( )φ ( )φ ( )φ

Similarly, recursive structures like (4) are excluded as a prosodic constituent of level n may not dominate a constituent of the same level n, i.e. an ι-phrase does not dominate an ι-phrase.

(4)

( )ι
( )ι ( )ι
( )φ ( )φ ( )φ

From the 1990s onwards evidence on prosodic phrasing suggest that the organization of prosodic constituents does not obey exhaustiveness constantly. Similarly, prosodic domains can be recursive as well (Ladd, 1996). These observations led to a less restrictive way of how prosodic constituents are hierarchically organized, formalized as violable constraints by Selkirk (1996).

Ideally, all prosodic constituents of the prosodic hierarchy in (1) are subject to characteristic phonological processes which would justify the incorporation of each constituent into a phonological model (Selkirk, 1984, 1990; Nespor & Vogel, 1986). In the literature, however, there is no agreement on the actual number and types of prosodic constituents that constitute the prosodic hierarchy, nor is there agreement on whether the prosodic constituents are strictly hierarchical. Most phonologists assume the Prosodic Word and the Intonational Phrase as constituents of the prosodic hierarchy. However, intervening prosodic domains are debated (Jun, 2005). Often, two levels of prosodic phrasing between the Prosodic Word and the Intonational Phrase are assumed. They are referred to as *Minor Phrase* and *Major Phrase* (Poser, 1984; Selkirk, 1986), *Accent Phrase* and *Intermediate Phrase* (Beckman & Pierrehumbert, 1986a; Gussenhoven, 2004; Jun, 2005) or *Clitic Group* and *Phonological Phrase* (Nespor & Vogel, 1986). For the present purpose, one level of prosodic phrasing between the Prosodic Word and the Intonational Phrase will be assumed, which, following Féry (2010b), I will call the Prosodic

Phrase (φ-phrase). This can be recursive and covers elements affected by processes that are attributed to the syntax-prosody mapping.

The notion of recursion is largely accepted in the more recent literature on prosodic phrasing and implies a set of prosodic domains which can be repeated at each level of the hierarchy. Wagner (2005); Selkirk (2009); Féry (2010c); Féry & Schubö (2011); Ladd (1996); Féry & Ishihara (2010); Ito & Mester (2009, 2011) give detailed overviews and motivations of recursive phrasing. On the other hand, Scheer has a contrasting view; he argues indisputably for non-recursivity in phonology (Scheer, 2012).

Have a look at (5c) which illustrates a recursive prosodic structure. The representation in (5b) describes a simple structure without recursiveness, and represents the syntactic structure in (5a) at the highest level of prosodic phrasing. By contrast, (5c) shows a recursive structure with embedded prosodic phrases, where the highest prosodic phrase P1, which is at the same time an Intonational Phrase, contains two embedded prosodic phrases (P2), of which the second P2 phrase, in turn, contains another embedded prosodic phrase (P3). Note, that the proposed structure in (5c) shows a better correspondence to the syntactic structure in (5a) than the one proposed in (5b).

(5) a. [$_{IP}$ [$_{NP}$ Anna] [$_{VP}$ hat [$_{NP}$ ein Eis] bestellt]]
b. [[ANNA]$_{P}$ [hat ein EIS bestellt]$_{P}$] $_{IP}$
c. [[ANNA]$_{P2}$ [hat [ein EIS]$_{P3}$ bestellt]$_{P2}$]$_{P1=IP}$

Besides the prosodic hierarchy as a relational based theory (Nespor & Vogel, 1986), a large number of other -competing- theories propose rules on the formation of syntax-related prosodic phrasing, among them, Edge-based theories (Selkirk, 1986), Optimality Theory and Alignment (Truckenbrodt, 1999; Selkirk, 2000; Féry & Samek-Lodovici, 2006), Match Theory (Selkirk, 2009, 2011; Elfner, 2012), Minimalism and Spellout (Ishihara, 2007; Kahnemuyipour, 2004; Kratzer & Selkirk, 2007). All approaches have the following in common: syntactic categories are represented as prosodic phrases, either by taking the syntactic constituent as a basis for mapping (i.e. Relational based theories and minimalism) or by considering the edges of syntactic categories as crucial (i.e. Alignment theories).

As shown below the Prosodic Phrase as well as the Intonational Phrase are intonationally defined and have relatively clear phonetic correlates as will be described in the next sections.

Prominence

An utterance can be divided into the next lower prosodic constituent and can thus consist of one or more Intonational Phrases (Selkirk, 1984). Intonational Phrases, and smaller prosodic units such as the φ-phrase, the Prosodic Word (ω), or the Foot (Ft) are the domains over which prosodic phenomena are expressed. Changing these prosodic phenomena/cues may induce a certain emphasis or accentuation on given words or phrases. Such emphasizing –

the assignment of accents – is strongly correlated with the degree of prominence a syllable carries within an utterance. Only prominent syllables are assigned accents (Pierrehumbert, 1980; Jacobs, 1993; Shattuck-Hufnagel & Turk, 1996). There are multiple phonological models that determine the prominence proportion between words (Halle & Vergnaud, 1987; Kager, 1995; Halle & Isardi, 1995). These phonological models have the concept of relative prominence in common, i.e. the prominence of a syllable is always defined relative to the prominence of other syllables in the same phrase.

Within the framework of the Metrical Phonology, different options are proposed with respect to the description of prominence proportions (Liberman & Prince, 1977; Halle & Vergnaud, 1987; Hayes, 1995; Halle & Isardi, 1995). The Theory of the Metrical Grid which was developed and put forward by Prince (1983) and Selkirk (1984) suggests that all words in the phrase are arranged along the bottom, and the rows of the grid indicate different levels of prominence. It is assumed that, in a first step, every syllable is associated with a beat. Then, a second beat is allocated to every stressed syllable. Finally, the most stressed syllable is assigned a third beat, as in (6). In this way, the metrical grid illustrates alternating prominent and non-prominent syllables by the number of their beats. A detailed description of prominence assignment in German can be found in Jacobs (1993).

(6)

```
                  x                             x
                  x                             x
 x     x   x      x   x             x     x   x      x  x
(die) kranken Schwestern           (die) Krankenschwestern
```

At the same time, the most prominent syllables of an utterance are so-called accent positions and therewith units which are assigned a specific prosodic marking. In German, such accent positions are realized by a change of the fundamental frequency on the accented syllable (Féry, 1993; Mayer, 1997). These accents are also called pitch accents (Pierrehumbert, 1980) and are associated with the most accented syllable within the sentence (Mayer, 1997). Such changes of the fundamental frequency are formulated in the framework of the autosegmental-metrical (AM) phonology in terms of the tone-sequence model by combining autosegmental and metrical phonology for tonal association with texts (Pierrehumbert, 1980; Ladd, 1996; Gussenhoven, 2004). Precursors of the AM intonation model can be found in the 1970s (Bruce, 1977), but it is Pierrehumbert's (1980) dissertation on American English intonation (see also Beckmann & Pierrehumbert (1986)) that is generally seen as the origin of this widely used framework. Since then, AM-models of intonation have been applied to various other languages (see Gussenhoven (1984) for Dutch; Sosa (1999) for Spanish). In the following, the general architecture and basic theoretical assumptions of the AM-framework are introduced, followed by a description of AM-models of German intonation as the framework of analysis for this thesis.

Theoretical Frameworks

The autosegmental-metrical framework

The AM-framework was developed to establish an abstract representation of intonation by describing intonational patterns as sequences of high (H) and low (L), occasionally also mid level (M) tones which determine the shape of the global pitch contour via transitions between the tones. This is why it is also referred to as a tone-sequence model. The tone-sequence model provided a grammar of intonation and formulated an algorithm for mapping outputs of this grammar into f0 contours (see Gussenhoven, 2004 for an overview). Secondly, the AM-framework represents prominence relations, modeled in terms of a metrical grid or tree structure (Liberman, 1975; Liberman & Prince, 1977), as illustrated above. Referring to the metrical aspect, Pierrehumbert (1980) suggests that tones can be classified as either pitch accents, boundary tones or phrase accents:

Pitch accents appear on lexically stressed syllables and are notated with the diacritic '*'. They can be either monotonal (H* or L*) or bitonal (H+L*, L*+H).[1] In the case of bitonal pitch accents, the starred tone indicates which of the two tones is associated with the accented syllable. The non-starred part of a bitonal pitch accent is called a leading tone, if it precedes the starred tone, and a trailing tone if it follows the starred tone. Originally, Pierrehumbert (1980) postulated a tonal inventory for English of two monotonal pitch accents (H* and L*), and five bitonal pitch accents (L*+H, L+H*, H*+L, H+L*, and H*+H), from which the bitonal pitch accent H*+H was removed in later accounts (Beckman & Pierrehumbert, 1986b; Pierrehumbert & Beckman, 1988).

Boundary tones are located either at the beginning or the end of a prosodic phrase and have demarcative function. In its origin, boundary tones were postulated to be associated with the edge of an Intonation Phrase (Pierrehumbert, 1980). They are marked with the diacritic '%' (Pierrehumbert, 1980; Féry, 1993; Grice & Baumann, 2002).

Phrase accents have a controversal theoretical status in intonational phonology (Gussenhoven, 2004). In the original version of the tone sequence model, Pierrehumbert (1980) postulated phrase accents (indicated with a minus symbol, L-) to capture the tonal movement from the last pitch accent towards the boundary tone of an Intonation Phrase. Later, that assumption was revised in that the phrase accent was attributed a demarcative function and served as boundary tone of the intermediate phrase (Pierrehumbert & Beckman, 1988). The status of the phrase accent, in particular its distinction from pitch accents and boundary tones are subject of much debate (see Gussenhoven, 2004, pp. 139-141). Also, models of German intonation

[1] The analysis of other languages include tritonal accents as well, see Gussenhoven (1984) for Dutch, Prieto et al. (2005) for Italian, or Kügler (2007) for certain German dialects.

posed different assumptions as to the presence or absence of the phrase accent (Féry (1993) as opposed to Grice & Baumann (2002)).

In the following, two AM-accounts on German intonation will be briefly summarized before turning to the difference that is made with respect to the phrase accent.

AM models of German intonation

AM approaches have been applied to German intonation by Wunderlich (1988); Uhmann (1991); Féry (1993); Grabe (1998); Grice & Baumann (2002); Peters (2006). The models mainly differ with respect to representational and tonal aspects, see Kügler (2007) for a detailed comparison. In the following, a description of the two widely used annotation models of German intonation is presented and distinctive assumptions on the derivation and representation of tones are summarized. The two models presented here are GToBI based on Grice & Baumann (2002) and Féry (1993).

German Tones and Break Indices (GToBI)

The first of these two models is GToBI, based on the autosegmental-metrical analysis of intonation, first mentioned in Grice et al. (1996); Reyelt et al. (1996) and fully described in Grice & Baumann (2002). Since then GToBI has been described in multiple publications and is seen as the standard system for annotating German intonation. In GToBI tonal patters are described on the basis of their surface appearance, i.e. annotated tones are not derived from an underlying abstract level of representation (see also Jun (2005) on that). Such a surface orientated approach allows for accents involving leading and trailing tones. As a result, GToBI provides an extensive accent inventory for the description of surface intonational patterns. Further, GToBI assumes two distinct levels of phrasing, the intermediate phrase (ip) and the Intonational Phrase (IP). Each phrase boundary is demarcated with its respective boundary tone, i.e. an intermediate phrase boundary is marked by a phrase accent, an Intonational Phrase boundary is marked by a boundary tone. An IP boundary always collapses with an ip boundary, and as a result, both boundary tone diacritics are combined in the GToBI notation if the phrase accent and the boundary tone carry different tonal values (i.e. a low phrase accent and a high Intonational Phrase boundary are labeled as L-H%). If the phrase accent and the boundary tone share the same tonal value, the boundary representation melt into one single label (i.e. a low phrase accent and a low Intonational Phrase boundary are labeled as L-%), see also Table 2.2 below. The following tonal inventory for modeling German intonation is proposed in Grice & Baumann (2002, 24)):

- two monotonal pitch accents: H*, L*
- four bitonal pitch accents: L+H*, L*+ H, H+L*, and H+!H*
- boundary tones: GToBI distinguishes boundary tones appearing at the end of an intermediate phrase from those occurring at the end of an Intonational Phrase. L-, H- , and !H- demarcate intermediate phrases boundaries and are marked with the diacritic '-'. L-

indicates a low, H- indicates a high, and !H- indicates a high but downstepped target point at the end of an intermediate phrase. L-%, H-%, L-H%, and H-^H% are used for transcribing Intonational Phrase boundaries.

GTobi has been revised in a comprehensive way by proposing to consult syntactic as well as semantic properties to account for the description of intonation (Féry, 2012).

German Intonational Patterns, Féry (1993)
The second model of German intonation is described in Féry (1993). Féry assumes two layers of tonal representation, an underlying and a surface layer. Pitch accents appearing on the surface layer are derived from underlying tones. Nuclear accents are always bitonal due to the theory of tonal linking (Gussenhoven, 1984) (see also Kügler (2007) for a summary on tonal linking as assumed in Féry (1993)), and are composed of a starred tone and (at least) one trailing tone (+L or +H). The trailing tone controls the melody from the nuclear accent towards the end of the Intonational Phrase. This analysis led Féry (1993) to refuse phrase accents in German; they are not necessary, neither to control the melody from the nuclear accent to the boundary tone, nor to delimit the intermediate phrase (in contrast to English (Pierrehumbert, 1980)). Rather, it is proposed that the unstarred part of the bitonal pitch accent, the tailing tone, captures both functions. For transcribing the most important German intonational patterns in the framework of the tone sequence model, the following tonal inventory is required (Féry, 1993: chapter 3):

- two pitch accents which can occur in nuclear as well as in prenuclear position: a simple fall (H*L) and a simple rise (L*H)
- three pitch accents which occur only in nuclear positions and which are restricted to a very specific contexts: a stylized contour (H*M), an early peak accent (HH*L), and a rise-fall (L*HL)
- one boundary tone: H% for transcribing rising IP boundaries; to model falling/ low IP boundaries the nuclear pitch accent is sufficient

Table 1.2 overviews the nuclear contours proposed in Féry (1993) and Grice & Baumann (2002), adopted from Grice & Baumann (2002, 24).

The status of the phrase accent
In the following, the crucial differences with respect to the presence or absence of the phrase accent is outlined and discussed. The existence of a phrase accent in German has been questioned in Féry (1993). Albeit Féry (1993) proposes distinct levels of prosodic phrasing, only the Intonational Phrase is demarcated with a boundary tone. The intermediate phrase boundary, in turn, is proposed be be tonally unspecified and only the spreading of the trailing tone until the end of the intermediate phrase characterizes the tonal movement from the pitch accent towards the phrase boundary. The concept of the phrase accent is denied in Grabe (1998) as well. Grabe (1998) examined German accents from a phonetic point of view by comparing

Nuclear contour		Féry	GToBI
Falling	1a	H*L	H*L-%
	1b		L+H*L-%
Rising-Falling	2	L*HL	L*+H L-%
Rising	3a	L*H	L*+(H) H-^H%
	3b		L*L-H%
	3c		(L+) H* H-^H%
Plateau	4	L*H	(L+) H* H-(%)
Falling-rising	5	H*L H%	(L+) H* L-H%
Early Peak	6a	H+H*L	H+!H* L-%
	6b		H+L* L-%
Stylized downstep	7	H*M	(L+)H* !H-%

Table 2.2: Summary of nuclear contours in German, adopted from Grice & Baumann (2002, 24)

the tonal phonology of English and German and questioned a phrase accent due "to the lack of evidence for the intermediate phrase in German" (Grabe, 1998, 47). Further, Grabe (1998) postulated tonal spreading as well to account for the pitch course between the accent and the boundary.

In contrast, GToBI assumes that an intermediate phrase boundary is tonally specified through a phrase accent (Benzmüller & Grice, 1998). This claim is based on experiments evaluating the low trailing tone in falling accents with respect to its alignment properties. The authors manipulated the amount of stressed and unstressed syllables after the nuclear accent, and observed that the alignment position of the low trailing tone is independent of the pitch accent but instead correlates with the location of the next stressed syllable, and take this observation as evidence for the presence of the trailing tone as independent phrase accent.

What remains to be clarified is the occurrence and relevance of the phrase accent in German. The mentioned studies investigated the nature of the phrasal accent phonetically and partly based their phonetic measurements on a very small database (Benzmüller & Grice, 1998), but none of the investigations provided a perceptual evidence for the existence of the phrase accent, which however, should be seen as a crucial foundation in favor of postulating phonological categories. Part of the perception results presented in chapter 4 suggest that there is no need to argue for a phrase accent as such. This issue will be picked up in chapter 4.

In the following, general assumptions on prosodic phrasing are overviewed, and section 2.2 summarizes literature on effects of prosodic phrasing on language comprehension to provide fundamentals for the concept of *prosodic cue weighting*, introduced in section 2.3 and relevant for the analysis and findings of the subsequent chapters.

Phrasing

The phonological phrasing of an utterance is very variable and thus cannot always be clearly defined by prosodic correlates (Cruttenden, 1986: 35ff; Féry, 1993: 59ff). Generally, utterances are divided into smaller phrases or constituents. Some of these constituents, such as the phonological phrase, and the Intonational Phrase exhibit systematic relations to syntactic constituents, although they have repeatedly been claimed not to be isomorphic to syntactic constituents. The rules relating syntactic and prosodic structure are known as syntax-prosody mapping. Syntax-prosody mapping concepts date back to Clements (1978), who explored tonal rules in Ewe and observed that multiple of these rules in Ewe do not spread leftwards across a syntactic boundary, contrary to the right boundary of a syntactic phrase. Similarly, Chen (1987) identified that in Xiamen Chinese the right boundary of a syntactic phrase seems to be the crucial place where tonal phenomena take place. Chen (1987) argued that the insertion of a tone group boundary coincides with each right boundary of a syntactic phrase (XP). In Xiamen, tone groups are identified by a tonal phenomena referred to as tone sandhi. Tone sandhi transforms all but the last tone in a tone group such that the appearance of an underlying tonal sequence within a tone group appears as T′ T′ T′ T (where T′ is the sandhi version of an underlying tone, and T is the underlying tone appearing unchanged on surface). Chen's (1987) observation on the phrasing of tone sandhi domains led to the generalization that the formation of prosodic phrases in Xiamen Chinese is formed by right boundary alignment with syntactic phrases; similar to Chi Mwini (Kisseberth & Abasheikh, 1974; Selkirk, 1986) which has also been claimed to show right boundary alignment of syntactic constituents (XP) with prosodic constituents.

Such grammatical restrictions of boundary alignment were formulated in Optimality-Theory (OT) frameworks by Prince & Smolensky (1993) in the format of ranked and violable constraints to describe phonological structure formation resulting from the syntax-prosody mapping. One type of constraints describe the relationship between syntactic structure and phonological structure in terms of alignment requirements between the edges of syntactic and prosodic constituents (Selkirk, 1986, 2000, 2005). So-called alignment constraints formulated as ALIGN-XP,R or ALIGN-XP,L demand that the left or the right edge of a syntactic constituents must coincide with the right or left edge of a prosodic constituents. Besides Selkirk's constraints on edge alignment, Truckenbrodt (1995, 1999) formulated another constraint named WRAP-XP, that equally determines the syntax-prosody relations on the level of the phonological phrase by forcing certain prosodic phrasings. WRAP-XP requires each XP to be contained in a phonological phrase. Later proposals formulated as Match theory (Selkirk, 2009, 2011; Elfner, 2012) force the mapping of syntactic constituents and prosodic ones by assuming three basic correspondence constraints, governing different levels of the prosodic hierarchy for either the syntax-prosody correspondence, or the prosody-syntax correspondence. Constraints governing the syntax-prosody relations call for correspondence between syntactic constituents in the syntactic representation (syntactic clause, syntactic phrase, syntactic word) and prosodic constituents in the prosodic representation (Intonational Phrase (ι), phonological phrase (φ), and the prosodic word (ω). This may be illustrated as in (7):

(7) **Syntax-prosody Match Constraints**:
MATCH-CLAUSE: Syntactic clause –> Intonational phrase (ι)
MATCH-PHRASE: Syntactic phrase (XP) –> Phonological Phrase (φ)
MATCH-WORD: Syntactic word –> Prosodic Word (ω)

Syntax-prosody match constraints are violated when there is a syntactic constituent in the input that has no corresponding prosodic constituent in the output. Similarly, prosody-syntax correspondence is governed by constraints which evaluate the relationship between prosodic constituents of the phonological representation and syntactic constituents in the syntactic representation. The correspondence constraints are shown in (8).

(8) **Prosody-syntax Match Constraints**:
MATCH-ι: Intonational phrase (ι) –> Syntactic clause
MATCH-φ: Phonological Phrase (φ) –> Syntactic phrase (XP)
MATCH-ω: Prosodic Word (ω) –> Syntactic word

The same logic as for the syntax-prosody match constraints is operative. Prosody-syntax match constraints are violated when a prosodic constituent in the output has no corresponding syntactic constituent in the input representation. Match Theory thus assumes a direct interaction between the various types of constraints that are responsible for determining prosodic constituency, requiring a syntactic category be exactly matched with a prosodic category. For a discussion and comparison of the Match Theory with Edge-based Theories, as well as for further discussions of the particular details of Match Theory, see the dissertation of Elfner (2012) who provided a detailed description of prosodic structure formation in Conamara Irish.

The next section shows that prosodic constituents resulting from the syntax-prosody mapping are prosodically demarcated by prosodic correlates. The most reliable prosodic correlates for German are boundary tones, lengthening of the syllables preceding the end of a prosodic phrase (prefinal lengthening), and pauses between prosodic phrases, but other correlates are possible as well.

Suprasegmental correlates of prosodic phrases

The following listing summarizes the most important prosodic correlates of Prosodic- and Intonational Phrases:

Boundary tones are local changes of the f0-movement at the beginning or end of an Intonational Phrase that do not express prominence (Cruttenden, 1986; Beckman & Pierrehumbert, 1986; Féry, 1993).

Prefinal lengthening: in many (typologically unrelated) languages the duration of the final syllable/ vowel preceding a prosodic phrase is increased (Vaissière, 1983; Wightman et al., 1992).

Pauses: a prosodic phrase can be preceded or followed by a pause; thereby the duration of the pause is dependent on the boundary strength, i.e. the pause duration increases with higher prosodic domains (Féry, 1993; Mayer, 1997; Grosjean & Collins, 1979; Fant & Kruckenberg, 1998; Horne et al., 1995; Fletcher, 2010).

Anacrusis: the production of syllables with an increased speaking rate and reduced articulatory preciseness may signal the beginning of an Intonational Phrase (Cruttenden, 1986).

Register resets: the register defines the *topline* and the *baseline* of the pitch range within a prosodic phrase of a particular speaker, and so determines the vertical excursion of the f0-movement. Within a phrase the register always remains stable (Cruttenden, 1986: 54; Ladd, 1996). Thus, a register reset may signal a prosodic phrase boundary.

Declination: *declination* or *downtrend* marks the continuous decline of the f0 within an utterance (Gussenhoven, 2004). This decline is not dependent on the tonal structure of the utterance (Ladd, 1984). The reset of the relevant f0 values are bound to a prosodic phrase boundary and may thus indicate phrasing (Liberman & Pierrehumbert, 1984; Jun, 2004; Ladd, 1988).

Each of the described prosodic features can apply separately, but when more than one of these features are present, the more likely it is to indicate a boundary. Moreover, boundaries of different strengths are marked differently, as revealed by the following studies.

A study by Terken & Collier (1991) shows that a professional speaker uses pauses, change of the f0-contour and prefinal lengthening to mark a strong boundary, i.e. a boundary between a Noun Phrase (NP) and a Verb Phrase (VP). The same speaker only uses a change of the f0-contour to mark a weak boundary, i.e. a boundary between a Noun and a Prepositional Phrase in a NP. Another study investigating the prosodic demarcation of phrasal boundaries was conducted by Strangert (1990) who reported that pauses are used more often to mark strong boundaries than to mark weak ones. A similar result for the feature preboundary lengthening is reported by Fernald & McRoberts (1996) who calculated data from Klatt (1976) and revealed that preboundary length is associated in 65% of the cases with a strong boundary. Additionally, pitch resets have been identified as being reliably correlated with syntactic boundaries (Jun, 2004; Ladd, 1988; Venditti et al., 1996). As an example, Ladd (1988) who investigated read speech reported that declination resets have been found to mark strong syntactic boundaries.

Likewise in Swedish, House (1990) and House et al. (1998) showed that listeners were most likely to interpret a declination reset as a strong syntactic boundary.

Apart from the prosodic correlates that make up a weak or a strong prosodic boundary, the tonal structure also offers important evidence with respect to the boundary strength, and thus the constitution of an intermediate/prosodic phrase or an Intonational phrase. In Féry's dissertation on German Intonational patterns (Féry, 1993), the tonal structure of a German Intonational Phrase has the following schematic representation (T* stands for a simple or complex pitch accent and T stands for an undefined tone of the value H or L):

(9) [[... T*T]ip [... T*T]ip (T%)]IP

An Intonational Phrase (IP) consists of one or more intermediate phrases (ip). Each ip contains at least one pitch accent. The limited inventory of postnuclear realizations in German does not require a phrase accent, as is the case in English where the phrase accent both, controls the tone from the final pitch accents to the right edge of the phrase, and delimits the intermediate phrase (Pierrehumbert, 1980; Pierrehumbert & Hirschberg, 1990).

The number and kinds of prosodic cues relevant for grouping ambiguous constituents into phrases is the central aspect in this thesis. It has been shown by several research papers that the above presented durational and tonal features seem to have a key role for listeners in the interpretation of ambiguous sentence structures (Collier et al., 1993; de Pijper & Sanderman, 1994; Steinhauer, 2011; Lehiste, 1973; Price et al., 1991; Beach, 1991). The results of these studies will be reviewed in the subsequent section 2.2 by summarising effects of prosodic phrasing on language comprehension.

2.2 Effects of prosodic phrasing on language comprehension

Studies that investigated the influence of prosodic information on language processing, in particular listeners' sensitivity to a prosodic boundary, have typically relied on three distinct experimental methods: (1) on perception studies using a scalar arrangement, or (2) on ERP studies investigating prosodic boundaries of different strength, and (3) on disambiguating studies. In the first line of studies, listeners are asked to make a scalar judgement about the strength of the boundary, or alternatively they are asked to decide whether the perceived unit, in writing, would either be marked by spaces, a comma or nothing at all (Collier et al., 1993; de Pijper & Sanderman, 1994). In the second type of studies, prosodic boundaries of varying strength are investigated by means of ERPs (Steinhauer, 2011). Steinhauer reported different responses to prosodic boundaries of different magnitudes. In the third line of experiments, the disambiguation studies, the listener is asked to judge the interpretation of ambiguous sentences or phrases such as *(one plus two) times three = 9*, as opposed to *one plus (two times three)*

= 7 (Lehiste, 1973; Price et al., 1991; Beach, 1991).

In the remaining part of the chapter, the above mentioned studies will be briefly described and their findings with respect to the role of prosody in the production and perception of structurally ambiguous sentences are presented.

Scalar judgement studies

With respect to the first line of studies, the scalar judgement studies, Collier et al. (1993) had listeners assign values of perceived boundary strength on a given scale ranging from 1 to 10. To accomplish this, 3 speakers read a set of 20 Dutch sentences like (10) that varied substantially in length and complexity to assure the occurrence of prosodic boundaries of varying strength.

(10) de intelligente onderzoeker reserveerde via de telefax
The intelligent researcher reserved via fax.

Subsequently, 19 listeners were asked to assign prosodic boundary strength values to each of the 175 word boundaries in the material. The calculated perceptual boundary strength for each word boundary resulted from the mean of the 19 given scores per word boundary. Then a correlation was established between the strength of the perceived boundaries and three prosodic variables: melodic discontinuity, declination reset and pause. The results are interesting in that they show that speakers differed in their strategies of marking prosodic boundaries, while listeners were very consistent in the perceptual weight they attributed to the prosodic cues.

ERP studies

In studies using Event-Related potentials (ERPs), different results for the processing of boundaries with varying strength were obtained. In a pioneering study, Steinhauer et al. (1999) recorded the on-line response to the processing of sentences containing either one or two prosodic phrase boundaries in sentences like *Peter verspricht ( ]IP) Anna ( ]IP) zu {arbeiten; entlasten} und das Büro zu putzen*. For the perception of each prosodic boundary, the so-called closure positive shift (CPS) has been manifested as a reliable and replicable ERP component. In a more recent ERP study, Steinhauer (2011) showed that the CPS varied as a function of boundary strength. He manipulated the strength of multiple competing boundaries in locally ambiguous garden path sentences like (11) and obtained different characteristics of the CPS for each boundary strength. The sentence material contained a subordinate clause with the verb *approaching* that can be intransitive or transitive, resulting in a temporary syntactic ambiguity between Early and Late Closure[2]. The beginning of the sentence through the phrase *the people* can continue as either Early Closure (closing the subordinate clause early at the intransitive verb *approaching* (11a)) or Late Closure (closing the subordinate clause late at the direct object *the people* (11b)).

(11) Whenever a bear was approaching #1 the people #2
a. (EC) ... would run away.
b. (LC) ... the dogs would run away.

42 EC/LC pairs with 2 prosodic boundaries each were recorded. The author then manipulated each boundary by adding 320 ms or 80 ms, to create IPh-compatible and intermediate size boundaries. Full permutation of boundary strengths resulted in 16 prosodic conditions (4 levels of boundary strength per boundary) for both EC and LC sentences. All prosodic boundaries elicited a CPS at the position of the boundary. However, the evoked component showed graded amplitudes and latencies that varied parametrically as a function of boundary strength and was largest for the strongest boundary (ι-boundary) and smallest (or absent) for the weakest boundary (φ boundary). The data strongly suggest a gradient CPS pattern reflecting the parametric manipulation of boundary strength.

Disambiguating studies

In a landmark study, Lehiste (1973) conducted a combined production and perception experiment for which she selected a set of 15 ambiguous English sentences which were recorded three times by four naive speakers. The first time, the speakers had to read the sentences and then were asked to indicate the meaning they had in mind. After pointing out the possible meanings to the speakers, the sentences were recorded a second and a third time asking the speaker to make a conscious effort to convey each of the sentences' meanings. All three productions of each sentence by each speaker were presented to listeners who were asked to identify the meaning intended by the speakers. The results indicated that 10 out of 15 sentences were disambiguated. The acoustic analysis revealed that word duration appeared to be the primary cue speakers used to disambiguate the sentences. The duration of words increased directly preceding stronger boundaries, i.e. the word *men* in (12a) precedes a stronger boundary due to the closure of the prosodic phrase *old men* and is thus lengthened as compared to *men* in (12b) which appears to be phrase initially (I follow the common convention and indicate prosodic boundaries with the pipe symbol "|" . The diacritic "||" stands for a strong boundary). Other correlates of strong boundaries found in Lehiste (1973) were laryngealization and insertion of pauses, but they happened to be less systematic.

[2] Early and Late Closure refer to universal structural parsing principles formulated in the Garden-path Model (Frazier & Fodor, 1978), see chapter 3 for a description, and for a detailed overview see Frazier (1987); Frazier & Rayner (1982).

(12) a. the (old men) (and woman) - the old | men || and woman
b. the old (men and woman) - the old || men | and woman

Similarly, Price et al. (1991) used a wide range of syntactic ambiguities and showed that listeners successfully disambiguate the spoken sentences. In a production task, four trained speakers read sentences in disambiguating contexts. In a listening task, participants were asked to identify from which context an ambiguous utterance was likely to have come. The results were consistent with Lehiste (1973) in that listeners were able to perform above chance, and they performed better on some sentences than on others based on their bracketing. Those utterances that were successfully disambiguated tended to involve Intonational phrase boundaries at major syntactic boundaries, and word durations increased in preboundary positions.

Equally, for the verification of a prosodic influence on sentence comprehension, Beach (1991) examined to what degree duration and f0 contributed to ambiguity resolution. The author examined sentence fragments like (13a) where the subject noun phrase is followed by the verb, and fragments like (13b) where the subject noun phrase is followed by the verb and the beginning part of the postverbal noun phrase.

(13) a. Jay believed ...
b. Jay believed the gossip ...

The critical ambiguous word, i.e. the main verb, was manipulated with respect to duration and f0 so that the sentence fragments were prosodically agreeing with either a direct object or clause complement continuation. In the comprehension task listeners were presented with sentence fragments up to the manipulated word and were asked to decide which of the two alternative complete sentences they thought the fragment was likely to have come from, i.e. the direct object or the sentence complement sentence. Independent of fragment length, small but significant effects were found when participants had to judge the congruence of sentence-initial prosody and intended continuation. The results suggest that prosody can be used in sentence interpretation before disambiguating syntactic-semantic cues are encountered.

The next section gives a survey of studies that have the individual impact of prosodic cues for disambiguation under scrutiny.

2.3 Prosodic cue weighting

In the aforementioned sections it has been shown that, in principle, three prosodic cues are strongly linked with phrasal boundaries. First, segments at the end of a phrase may be longer than in the middle of a phrase (Beckman & Edwards, 1990; Cooper & Paccia-Cooper, 1980; Klatt, 1975; Wightman et al., 1992; Vaissière, 1983). Second, f0 may signal a phrase boundary through pitch reset or through a change of the intonation contour (Beckman & Pierrehumbert, 1986b; Pierrehumbert & Beckman, 1988). And thirdly, pauses between phrases may be longer than pauses within a phrase (Cooper & Paccia-Cooper, 1980; Scott, 1982; Féry, 1993; Mayer, 1997; Fant & Kruckenberg, 1998). The following sections summarizes papers revealing that these cues are used differently across languages. In particular, speakers of one language may rely on a subset of these cues more than speakers of another language when demarcating phrasal boundaries. Additionally, the presented papers suggest that the weighting of prosodic cues change over the time of development (Nittrouer & Miller, 1997; Nittrouer et al., 1993).

For example, the findings of Seidl & Cristià (2008) on the use of prosodic cues suggest that 4-month-old English infants rely heavily on all three prosodic cues while 6-month-old English infants only need the combined information of pitch and pause, or pitch and preboundary length to segment clauses successfully (Seidl, 2007). With respect to adult English listeners, multiple studies indicate that, when asked to interpret a boundary strength, pause is not a necessary cue, but f0 and prefinal lengthening are both important (Aasland & Baum, 2003; Kjelgaard & Speer, 1994; Scott, 1982; Streeter, 1978; Wightman et al., 1992).

For instance, Aasland & Baum (2003) consecutively changed the prosodic cues accompanying prosodic boundaries by increasing or compressing vowel and pause duration within the phrase *pink and black and green*. Also, Scott (1982) investigated whether listeners use the feature *prefinal lengthening* to parse syntactically ambiguous sentences, such as *Kate or Pat and Tony will come*, where the phrase boundary after *Kate* represents one meaning, and after *Pat* another meaning. Streeter (1978) examined the individual influence of amplitude, pitch and duration on phrase boundary perception in ambiguous algebraic expressions, such as *(A plus E) times O* and *A plus (E times O)* while varying the suprasegmental features, respectively. The results of these studies were consistent in showing that listeners were better able to interpret the ambiguous constructions when prefinal lengthening and pitch were correlated with a prosodic boundary than when pause changes were. Similarly, Lehiste et al. (1976) point out that in terms of cue use in perceiving syntactic boundaries in ambiguous constructions, prefinal length is heavily weighted for American English native speakers. To sum up, the above studies indicate that for English listeners, pause does not appear to be a reliable marker of syntactic boundaries, while the feature prefinal lengthening is, see also Martin (1970a) and Henderson & Nelms (1980).

With respect to Dutch learners, a somewhat different picture emerges in that they seem to be more reliant on pauses to detect clause boundaries (Johnson & Seidl, 2008). In a study on Dutch, Johnson & Seidl (2008) investigated infants' ability to extract clauses like *koude pizza smaakt niet zo goed* (*cold pizza doesn't taste so good*) from speech and found that 6-month-old Dutch infants can segment the speech signal into clauses but are unable to do so when the pauses marking clause boundaries are removed. Intonation differences between Dutch and English provide a possible explanation. Dutch has a narrower pitch range than English (Collins & Mees, 1981; Willems, 1982) which may make listeners more dependent on pause information when segmenting clauses. Additional support for Johnson and Seidls' (2008) results are provided by Sanderman & Collier (1997) who suggest that adult Dutch listeners seem to rely heavily on pause cues to detect clause boundaries.

When looking at Hindi, a 'phrase language' (Féry, 2010a; Patil et al., 2008), in which the melody of sentences is primarily determined by invariant phrasal tones, a completely different use of prosodic cues for demarcating phrasal boundaries in production is observable. In a study by Féry & Kentner (2010) the prosodic renderings of simple as well as embedded groupings of three- and four-name conditions were studied. While the equivalent study for German showed speakers' capacity to reflect each syntactic grouping condition by prosodic means, Hindi speakers did not contrast the different groupings reliably, neither by pitch nor by durational differences. Féry & Schubö (2011) report similar results on center-embedded relative clauses. The authors propose that the difference between German and Hindi is due to their different intonational systems, with German being an Intonational language, using distinct pitch accents to convey different meanings, and Hindi being a Phrase language, and thus, being much more rigid in its use of intonation.

The above cited studies have contributed important insights: they clearly show that the prosodic reflection of syntactic structure in production and the prosodic cues used for disambiguation in perception do not show a universal strategy, nor is there agreement with respect to the individual prosodic cues that support disambiguation. Instead, the language-specific role of prosody is crucially important when ascertaining both, whether and how syntactic relations are reflected in prosodic structure.

2.4 Motivation and aims

The above literature review has shown that models of German intonation generally include different levels of prosodic phrasing, among them the Intonational Phrase and a smaller constituent, here defined as Prosodic Phrase (φ-phrase). Intonational and φ-phrases usually appear to be cued through language specific prosodic correlates. Furthermore, it has been shown that listeners pay more attention to certain prosodic correlates than others (or weight them differently) in their perception of suprasegmental units, see Aasland & Baum (2003); Kjelgaard & Speer (1994); Scott (1982); Streeter (1978); Wightman et al. (1992) for English, or Sanderman & Collier (1997); Johnson & Seidl (2008) for Dutch. However, surprisingly little effort has been spent on the question as to whether the sum of all prosodic cues are necessary for auditory sentence disambiguation in German, or whether each prosodic cue is equally efficient in disambiguating sentences. Therefore, information on German cue weighting is sparse, especially with respect to the use of prosodic cues in different boundary environments.

Specifically, the experiments of the present thesis are designed to extent the concept of prosodic cue weighting for German, and are intended to provide new evidence for ambiguity resolution in German by comparing listeners' ability to disambiguate sentences carrying different amount of prosodic information. The experiments of this dissertation sequentially neutralize each prosodic cue (*prefinal lengthening, pause duration, and f0*) of sentences containing either φ-phrase or ι-phrase boundaries to investigate whether listeners' performance on a sentence completion task remains the same or is reduced as the result of the loss of each correlate (cf. chapter 4 and 5). Moreover, the experiments were designed to take into account one further factor that can be expected to influence prosodic cue weighting in German, namely the boundary strength, to test whether the level of prosodic phrasing interacts with acoustic cues listeners use for ambiguity resolution. As will be shown, this will not only allow us to assess previous claims and create novel predictions on prosodic cue weighting in German, but it will also support models of speech perception that predict prosody to play a role as input structure for the parser, and thus constraints syntactic parsing (Marcus & Hindle, 1990; Speer et al., 1996; Schafer, 1997). As a consequence, the syntactic structure of a sentence is generated on the basis of its prosodic structure. Based on the experimental results on prosodic cue weighting, we propose Optimality-Theoretic models that capture the variation of prosodic boundary correlates relevant for prosodic boundary formation in speech perception (cf. chapter 6).

3

General aspects of design and stimuli

As has been demonstrated in chapter 2, numerous studies examined the influence of prosody on ambiguity resolution. Far fewer studies examined the impact of each individual prosodic cue on the processing of ambiguous structures, a concept referred to as *prosodic cue weighting* in the literature. This thesis will explore prosodic cue weighting in German by investigating ambiguous sentence structures containing two successive NPs, in which the second NP can either be interpreted as a possessive modifier of the first NP (genitive-reading), or as a verbal argument (dative-reading), resulting in a φ-boundary after the first NP only in the dative reading, but not in the genitive-reading.

This chapter provides an overview of the experimental design and discusses some general considerations on the auditory stimuli used for the experiments conducted within this thesis. In particular, section 3.1 introduces the conditions of the experiments, and section 3.2 is concerned with the general impacts of structural complexity on parsing. Specifically, syntactic considerations on the experimental stimuli, a rating study and a corpus based frequency analysis are discussed. The conglomerate of these analyses on German ambiguous NPs point to an important difference between these two sentence structures in the context of all experiments presented in chapter 4 and 5, which must be taken into consideration when interpreting the results, namely that there is a preference for one of the two sentence structures on the basis of the syntactic preference for the dative structure, whereas there is no such preference for the genitive structure. In this way, one should expect to find initial differences in the dynamics of processing between the two conditions which cannot be attributed to the prosodic manipulation. The prosodic manipulation then adds information to the structure and either facilitates or complicates disambiguation, depending on the importance and magnitude of the prosodic cue, respectively.

3.1 Properties of the stimulus material

Two lines of experiments are conducted. Each experimental line consists of a production experiment, intended to identify the prosodic cues used by speakers in disambiguating sentence structures, and a series of perception experiments examining the individual impact of each prosodic cue on listeners disambiguation. The experiments subsumed in chapter 4 focuses on the role of φ-boundaries and their corresponding prosodic correlates on listeners' disambiguation of syntactically ambiguous sentences. The second line of experiments, described in chapter 5, examines the role of Intonational Phrase boundaries on listeners' disambiguation. Through these two series of experiments, the disambiguating role played by φ-boundaries, as well as the role of larger ι-boundaries and their individual prosodic influence on sentence comprehension are under examination.

All perception experiments use the same forced-choice paradigm task. Participants listen to sentence fragments that contain a structurally ambiguous NP. At the offset of each sound file, participants are asked to complete the auditory sentence fragments by choosing one of two disambiguating sentence continuations offered in a questionnaire. For the experiments involving φ-boundaries (cf. chapter 4), verbs served as disambiguating sentence material; for the experiments involving ι-boundaries (cf. chapter 5), either Prepositional Phrases (PPs) (in most of the cases), or ditransitive verbs served as disambiguating sentence material. The disambiguating verbs used for the φ-boundary experiments are either transitive or ditransitive, and therefore, disambiguate the fragments either towards the genitive structure (intransitive verb), or the dative structure (ditransitive verb). For the ι-boundary experiments, PPs disambiguated the sentence fragments towards the genitive structure, whereas the offered verbs disambiguated the structures towards the dative interpretation. The auditory fragments for all perception experiments incorporate two independent variables: (i) the location of a phrase boundary, and (ii) the boundary type. Table 3.1 summarizes the experimental conditions for all experiments.

Experiment	Condition	Auditory Fragment	Visual text
φ-boundary	GEN	(NP1 (NP2)) (NP3)	trans/ditrans verb
	DAT	(NP1) (NP2) (NP3)	trans/ditrans verb
ι-boundary	GEN	Verb (NP1 (NP2))	PP/ditrans verb
	DAT	Verb (NP1) (NP2)	PP/ditrans verb

Table 3.1: Summary of conditions for all experiments

3.2 Impact of complexity on parsing

Processing theories are interested in the question of whether one of two ambiguous readings is preferred in the process of language comprehension. For example, it has been claimed that a reading is preferred when it is syntactically simpler in that it contains fewer syntactic nodes (cf.

Garden-Path-Model with its parsing principle Minimal attachment (Frazier & Fodor, 1978)). In the following section such syntactic considerations and their implications for the ambiguity used in chapter 4 and 5 are presented.

Syntactic account

In psycholinguistic research, numerous studies on language processing investigate local or global ambiguities. According to an early and influential approach, the Garden-path Model (Frazier & Fodor, 1978), ambiguity resolution is explained by proposing a serial parser operating on a set of universal structural principles, referred to as *Minimal Attachment* (MA) and *Late Closure* (LC). One of the proposed parsing principles, MA (Frazier & Fodor, 1978), predicts that the parser attempts to interpret sentences in terms of the simplest syntactic structure consistent with the input that is known at the moment. However, when two possible readings of an ambiguous sentence are similar in syntactic complexity, the second parsing principle LC, operates, and drives the parser to attach new incoming sentence material to the most recent node, i.e. to the clause or phrase which is currently processed (for a detailed overview see Frazier (1987); Frazier & Rayner (1982); De Vincenzi & Job (1993, 1995); for Recency see also Gibson et al. (1996)). The principle of MA is illustrated by the locally ambiguous sentence in (14), in which the NP *the answer* can either be interpreted as the complement of the verb (14a) or as the subject of the complement clause (14b). MA predicts a preference for the complement reading analysis (14a) which is claimed to be less complex and therefore preferred.

(14) a. The girl IP[VP[V[knew NP[the answer] by heart]]].
(14) b. The girl IP[VP[V[knew CP[IP[NP[the answer VP[was wrong]]]]]]].

example taken from Rayner & Pollatsek (1989)

As one incrementally process the sentence, the MA principle produces a garden-path effect on the disambiguating region *was* in (14b), because, for the correct reading, the initially assigned complement reading needs to be changed to the more complex subordinate clause reading. Experimental data evidenced that this change can be empirically measured. Grammaticality judgments (Schütze, 1996), for example, showed that reading times were significantly longer for sentences whose initial interpretation was not in accord with the MA strategy, as opposed to sentences in which the preferred reading conforms with the MA principle. Similarly, eye-tracking studies indicated an increase in processing time at the disambiguating position *was* (Ferreira & Henderson, 1990), which could be taken as further evidence for revealing a garden-path effect in (14b). The MA principle can be successfully applied to a large number of constructions in many languages (for a detailed overview, see Frazier & Clifton (1996)), it however, can by no means explain all parsing preferences in cases of syntactic ambiguity. Additionally, the parsing principles proposed by the Garden-path Model have come to be challenged over time from constraint-based models (McRae et al., 1998a) who drop the syntax-first assumption,

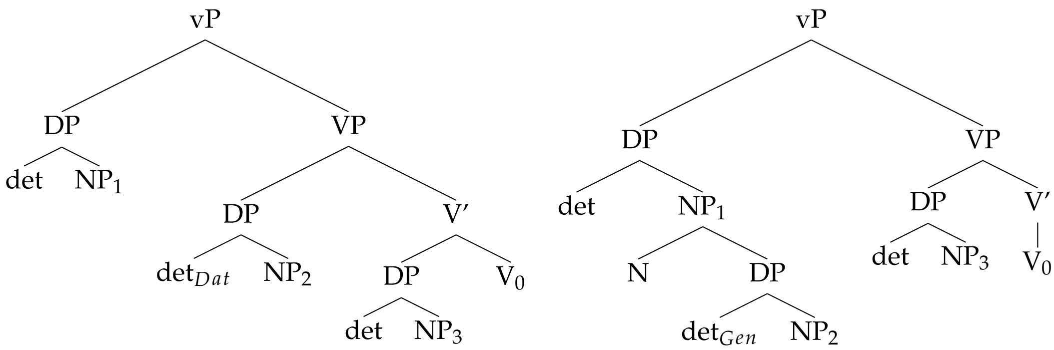

Figure 3.1: Structural representation of the dative-reading on the left. Example sentence of the dative-reading: *Neulich hat der GärtnerNP1 der ReiterinNP2 den BaumNP3 gezeigt, der morsch war.* 'Recently the gardener showed the horsewoman a tree, that was decaying.' Structural representation of the genitive-reading on the right. Example sentence of the genitive-reading: *Neulich hat der GärtnerNP1 der ReiterinNP2 den BaumNP3 gefällt, der morsch war.* 'Recently the gardener of the horsewoman chopped the tree, that was decaying.'

and instead allow semantic, lexical, plausibility information etc. to play an equal role during ambiguity resolution.

With respect to the experimental sentences used in chapter 4, (cf. example 15), there is an ambiguity between the interpretation of the second NP, which can either modify the first NP (Figure 3.1, right), or be a dative object (Figure 3.1, left). According to the proposed parsing principle MA, the parser would prefer the dative-reading since that structure is simpler and its ambiguous NP2 is less embedded compared to the genitive-reading which contains a deeper embedded ambiguous NP (Chomsky, 1993, 1995, 2000). Strictly speaking, in the genitive-reading, *Gärtner* does not represent a proper NP but a noun (N). The actual NP is represented by *der Gärtner der Reiterin*. However, for the sake of the following analyses and for reasons of comparability to the dative structure, we use NP as a general label for N, independent whether or not if forms a NP.

Rating study

The preference for the predicted dative-reading has been corroborated in a sentence rating study. Each ambiguous sequence including both possible disambiguating sentence continuations were printed separately on A4 (see Table 3.2). The subject's task was to read the sentences thoroughly and then select one of the two sentence continuation which they thought would be the appropriate continuation. 48 students from the University of Potsdam, all naïve to the purpose of the experiment, participated in the rating study for course credit. To avoid order effects, 24 subjects (group 1) were given the disambiguating dative verb on top of the genitive verb, the other 24 subjects (group 2) were given the reversed ordering of the two sentence continuations.

group	sentence	condition	verb information
1	Neulich hat der Gärtner der Reiterin den Baum	DAT	gezeigt, der ...
		GEN	gefällt, der ...
2	Neulich hat der Gärtner der Reiterin den Baum	GEN	gefällt, der ...
		DAT	gezeigt, der ...

Table 3.2: Presentation of the experimental stimuli in the rating study for group 1 and 2. Sentence *Neulich hat der Gärtner der Reiterin den Baum {gezeigt, gefällt}, der morsch war.* 'Recently the gardener showed the horsewoman a tree.' / 'Recently the gardener of the horsewoman chopped a tree.'

Results

Independently of the presented order, the dative verb was selected more often than the genitive verb. Specifically, in group 1, the dative verb was chosen 13 times, the genitive verb 11 times. For group 2, the dative verb was chosen 16 times, and the genitive verb 8 times. The ratings were evaluated with a Fisher's Exact Test for Count Data and revealed no significant effect (p=0.55) of grouping for the verb selection, thus indicating no difference in terms of a preference of a sentence structure that is systematically attributable to the order of appearance of the sentence continuation. However, the odds ratio given in Table 3.3 indicate that it is 1.52 times more likely to select the dative sentence continuation over the genitive sentence continuation in the present rating study. In general, in logistic regression analyses, the odds ratio expresses how many times more likely it is that a certain event occurs than another event.

group	p-value	odds ratio
1 and 2	0.55	1.52

Table 3.3: Results of the Fisher exact test

Excluding frequency effects

In order to exclude the possibility that a preference for one particular sentence structure is only a function of sentential frequency, we did a corpus analysis using the TIGER corpus[1],

[1] The TIGER-corpus (Brants et al., 2002, 2004) was developed within the framework of the DFG-funded TIGER-Project (1999-2004) to generate a syntactically annotated corpus implemented with a search engine. This endeavor had been carried out from the institutes of the Universities Saarbrücken, Stuttgart and Potsdam. The corpus which is based on texts from the German newspaper Frankfurter Rundschau contains 50.000 syntactically annotated sentences (i.e. about 900.000 tokens). Syntactic information (phrase structure and grammatical functions) as well as part of speech,

which presents a very useful tool to investigate the distribution and frequencies of sentence structures. For the genitive and dative condition, a sequence of three successive NPs was requested. For each condition request, the NPs were individually labelled for case and had to be followed by a verb, i.e. genitive condition request: NPnom . NPgen embedded . NPacc followed by the verb; dative condition request: NPnom . NPdat . NPacc followed by the verb. Appendix B lists the entire corpus request for both sentence structures. The request for the genitive condition resulted in 13 hits, the one for the dative condition yielded 14 congruities. Therefore, the results of the corpus analysis showed an even distribution of the frequencies for both structures (cf. Table 3.4).

condition	query-sequence of NPs	output
GEN	NOM-GEN-ACC-Verb	13
DAT	NOM-DAT-ACC-Verb	14

Table 3.4: Results of the corpus inquiry - frequency of occurrence

These results suggest that a possible bias towards the dative structure may be attributed to a difference in syntactic complexity (according to the parsing principle MA) rather than a result of a difference in occurrence.

ERP evidence

Parsing preferences for ambiguous NPs in German have been also reported by Augurzky & Schlesewsky (2010). In an ERP study, they examined the interplay between prosodic phrasing and case information of German NPs. NP1-NP2-V constructions were investigated, in which the second NP can either be interpreted as a possessive modifier of the first NP, or as a verbal argument. The ERP data show an N400 at the onset of the intransitive verb for the genitive-reading, reflecting the increased processing cost for the clause-final integration of the possessive modifier (adjunct) as opposed to the argument. This result is in accordance with previous studies on argument-adjunct ambiguities (Clifton et al., 1991; Schütze & Gibson, 1999) and suggests in our case a general advantage for the dative-reading over the genitive-reading.

The excursion on the preference of the experimental stimuli invariably suggests that the choice of the syntactic computation in an ambiguous construction such as NP1-NP2-NP3-Verb is dependent on the syntactic complexity. The prediction of the parsing principle MA, the results of the rating study and the findings on the argument vs. adjunct preference (Augurzky & Schlesewsky, 2010) give evidence and justify the conclusion that syntactic constructions containing an argument are favored over constructions that include an adjunct. For the experimental

morphology and lemma information is annotated. The distinction between arguments and adjuncts, for instance, is not expressed in the constituent structure, but is instead encoded by means of syntactic functions (i.e. dative vs. genitive).

sentences used in this thesis, this preference is unaffected by the frequency, which is evenly distributed among both structures, according to the results of the TIGER-corpus.

As a result of the above presented studies, an initial preference for the dative-structure is expected, independently of the performed prosodic manipulation in order to establish a prosodic cue weighting. The experiments allowing to formulate a prosodic cue weighting on φ-phrase boundaries are presented in the following chapter.

4

Prosodic cue weighting of φ-Phrases

Previous work has shown that speakers and listeners efficiently exploit prosodic information to make the meaning of syntactically ambiguous sentences explicit. However, quantifiable phonetic properties of prosody in speech production (*prefinal lengthening, pause duration and f0*) stand in a complex relationship to the percept they induce in the auditory domain. Not all measurable prosodic differences are actually used in sentence parsing (Lehiste et al., 1976; Streeter, 1978; Mo, 2008; Aasland & Baum, 2003; Kjelgaard & Speer, 1994; Scott, 1982; Wightman et al., 1992). This chapter presents a production experiment designed to examine speakers' use of prosodic cues. Speakers read out locally ambiguous sentences containing a sequence of three NPs with either a φ-phrase boundary or no boundary between the first and second NP. In a series of perception experiments, the prosodic cues were manipulated to verify which prosodic cue is used to disambiguate the structures in perception. The combined results of the production and perception experiments reveal a divergence of prosodic cues used by speakers and those used by listeners, as speakers exploit *prefinal lengthening, pause duration and f0* to disambiguate the structures, while listeners only need *prefinal lengthening* to identify the reading speakers assigned to the structures. Exclusive *f0 information* or exclusive information of *pause duration* do not allow listeners to disambiguate the structures.

4.1 Production experiment

Generally, prosody is known to influence syntactic analysis in ambiguous sentences (Marslen-Wilson et al., 1992; Kjelgaard & Speer, 1999; Millotte et al., 2008). Although there is ample evidence on the disambiguating role played by Intonational Phrase boundaries, the role of lower phrase boundaries appears to be an open question cross-linguistically (Price et al., 1991; Kjelgaard & Speer, 1999; Li & Yang, 2009). In order to scrutinize the prosodic cues exploited by speakers to prosodically mark lower phrase boundaries, a production experiment was conducted using locally ambiguous structures containing lower φ-phrase boundaries as presented in (15). These locally ambiguous structures represent appropriate sentence material to investigate the phonetic details associated with φ-phrase boundaries. Juxtaposing both sentence conditions (13a and 13b) makes the deflection of prosodic cues transparent (cf. Figure 4.1 and Figure 4.3).

Material

The experimental sentences contain a matrix and subordinate clause. The matrix clause contains a sentence adverb followed by three NPs and a verb. In the matrix clause there was a local ambiguity with respect to the second NP, which could either be interpreted as a possessive modifier of NP1 (15a), or as a verbal argument (15b). In both readings, the linear succession of words up to and including NP3 was identical and the sentences were disambiguated on encountering the verb. A transitive verb disambiguated the genitive-reading (15a), while the dative-reading was disambiguated by a ditransitive verb (15b).

(15a) genitive-reading

Neulich hat der Bruder der Nachbarin ein Haus gesehen, das zerfallen war.
recently the brotherNOM the neighbourGEN a houseACC see, that was decaying
Recently the brother of the neighbour saw a house, that was decaying.

(15b) dative-reading

Neulich hat der Bruder der Nachbarin ein Haus geschenkt, das zerfallen war.
recently the brotherNOM the neighbourDAT a houseACC give, that was decaying
Recently the brother gave the neighbour a house, that was decaying.

The stimuli were controlled by number of syllables, stress pattern and gender. All NP1s were masculine disyllabic trochees, all NP2s were feminine trisyllabic trochees and all NP3s were of neuter gender and monosyllabic. The sentence material was highly sonorant to allow for a maximal accurate pitch analysis. 10 sentences per case condition were created. The resulting 20 sentences were interspersed with numerous fillers and fed into linger software. The experimental sentences were pseudo-randomized for each subject so that sentences of the same condition did not appear adjacent to one another and corresponding sentences of the two conditions had a maximal distance. Appendix A contains a complete list of all sentences used for the production study.

Participants

18 speakers participated in the experiment. All were female undergraduate students at the University in Potsdam and were residents of the surrounding areas. All were native speakers of German and reported no speech or hearing impairment. They either received course credit or were paid for participation.

Method

For each sentence, a context question in broad focus, spoken by a male native voice, had been previously recorded. The contexts were presented together with a target sentence both visually on screen and auditory over headphones. The items were presented on a 15 zoll computer screen. Participants were asked to read and listen to the context and then speak out the answer displayed on the screen as a response to the question. Subjects were familiarized with the task through written and verbal instructions, followed by three practice trials. In case of hesitations or false starts, participants were asked to repeat the sentence. Recordings took place in a sound-proof chamber equipped with an AT4033a audio-technica studio microphone, using a C-Media Wave soundcard at a sampling rate of 44.1 kHz with 16 bit resolution. Presentation flow was controlled by the experimenter, and participants were allowed to take a break whenever they wanted.

The target sentences were hand-annotated at the level of the constituent (see Table 4.1) and subjected to phonetic analysis using Praat software (Boersma & Weenink, 2013). The duration of the adverbial phrase, of each NP, plus the pauses between NP1 and NP2 and between NP2 and NP3 were measured. Pitch analysis was conducted using a Hanning window of 0.4 seconds length with a default 10 ms analysis frame. The pitch contour was smoothed using the smoothing algorithm (frequency band 10 Hz) to diminish microprosodic perturbations (Boersma & Weenink, 2013). Stylized pitch tracks were calculated. For this purpose, each constituent (see Table 4.1) was divided into five equal-sized intervals, and the mean F0 (in Hz) per interval was aggregated over all speakers and sentences for each interval. The resultant values were interpolated for each condition.

ADV	NP1	P1	NP2	P2	NP3
|Neulich hat|	|der Bruder|	#1	|der Nachbarin|	#2	|ein Haus|
recently has	the brother	#1	the neighbour	#2	a house

Table 4.1: Illustration of the labeling positions, i.e the adverbial phrase (ADV), the first NP (NP1), the pause between NP1 and NP2 (P1), the second NP (NP2), the pause between NP2 and NP3 (P2), and the third NP (NP3).

Results

A phonetic validation of the production data on the ambiguous region is required in order to gauge the prosodic cues used by speakers from those phonetic values used by listeners for disambiguation. *Prefinal lengthening, pause durations and f0 (maximum)* values on the critical ambiguous sequence in the genitive reading were compared with realizations of the dative reading. The following section gives the results of the phonetic evaluation. First, the duration results are reported. Second, the evaluation of the tonal contour, with a focus on the scaling of the pitch accents, are given. The results of the phonetic evaluation is indicative with respect to the phrasal structure which is presented in the last section.

Duration

This section examines word duration, testing the prediction based on previous research that the duration of words were longer when occurring in prefinal positions, i.e. preceding phrase boundaries. The mean duration of words in both sentence conditions is shown in Figure 4.1, where dark grey bars depict the mean durations of the words in the genitive condition and light grey bars those of the dative condition. For both experimental conditions, Figure 4.1 illustrates the averaged duration of the adverbial phrase (ADV), the first noun phrase (NP1), the second noun phrase (NP2), and the third noun phrase (NP3). The average duration time is presented in seconds (s) displayed at the y-axis. As for the dative condition, words of the first noun phrase (NP1) had a longer duration than those of the genitive condition within the same interval. The duration of words of the second noun phrase (NP2) were longer for the genitive condition compared to the dative condition. The duration of words of the third noun phrase (NP3) appeared to be longer for the genitive condition as opposed to the duration of words of the third noun phrase in the dative condition, as illustrated in Figure 4.1. Linear mixed-effects models confirmed the phonetic difference of the word duration for the first noun phrase (NP1), the second noun phrase (NP2), and the third noun phrase (NP3), see Table 4.2.

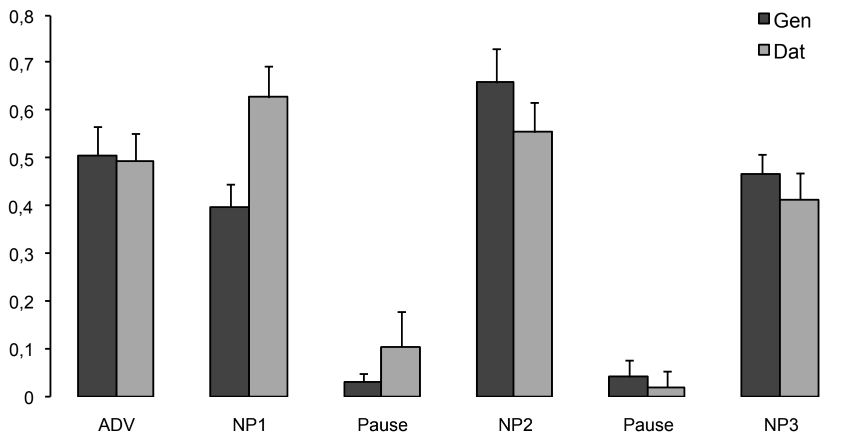

Figure 4.1: Mean duration times in s per constituent and pause across all speakers

Duration of pauses

The duration of pauses is hypothesized to differ between both sentence conditions as a function of signaling a different phrasal structure. Testing that prediction, the distribution and duration of pauses were evaluated with a praat scrip algorhythm to labelling a pause by an interval marking the start and the end of that pause on the basis of an intensity analysis. The following criteria defined a pause within a sentence: positive minimum duration (seconds)

Constituent	Comparison	Means (duration (ms))	t-scores
ADV	GEN vs DAT	504 494	0.87
NP1	GEN vs DAT	397 627	5.97*
NP2	GEN vs DAT	658 554	-4.21*
NP3	GEN vs DAT	465 413	-3.81*

Table 4.2: Word duration on the first four constituents of genitive and dative sentences as well as their statistical comparisons by means of t-tests. Absolute t-values above |2| are significant and are marked with an asterisk.

= 0.01, positive maximum intensity (dB) = 55. The results of the pause analysis revealed a difference of pause duration but not one of pause distribution. Two pause positions were identified. The distribution of these pauses were found to be located between NP1 and NP2, and between NP2 and NP3. The mean pause duration in both sentence conditions is shown in Figure 4.1; dark grey bars depict the mean pause durations of the genitive condition and light grey bars those of the dative condition. As for the dative condition, pauses between NP1 and NP2 showed a longer duration compared to the pause duration of the genitive condition. Pause duration between NP2 and NP3 for both conditions appeared to be marginally different, as illustrated in Figure 4.1. The corresponding statistical analysis using a linear mixed-effects model revealed a significant difference of the pause duration between NP1 and NP2 (Pause 1) but not one between NP2 and NP3 (Pause 2) for both sentence conditions, see Table 4.3.

Constituent	Comparison	Means (duration (ms))	t-scores
Pause 1	GEN vs DAT	28 104	2.94*
Pause 2	GEN vs DAT	40 18	1.37

Table 4.3: Duration of pause 1 and 2 of genitive and dative sentences as well as their statistical comparisons by means of t-tests.

Tonal contour

This section presents the results of the analysis of the tonal contour for each sentence condition, aiming to evaluate whether the overall tonal contour is influenced by a difference in argument structure. First, it reports the type and presence of pitch accents for each defined interval in both conditions (see Table 4.1. above), illustrated with an example pitch contour. Second, it discusses the scaling of the pitch accents and third, a summary of the phrasal structure for both conditions is offered that is based on the duration results as well as the pitch accent scaling.

A mean pitch contour, averaged over all 18 speakers was generated using a praat script that subdivided each defined interval (see Table 4.1) into five equal-sized subintervals. The mean F0 (in Hz) per interval was aggregated over all speakers and sentences for each interval. The resultant values were interpolated for each condition. The averaged pitch contour is depicted in Figure 4.3, where a solid line depicts the mean pitch contour in the genitive condition and a dashed line those of the dative condition. The overall pitch contour in both conditions is identical in showing rising pitch accents (L*H) on the adverbial phrase, on NP1, NP2 and

NP3. There is no tonal contrast, such as different kinds of pitch accents. Instead, there is only one relevant difference resulting from the scaling of the pith accents on NP1, NP2 and NP3 itself. For the dative condition, the mean pitch values of NP1, NP2 and NP3 are successively lowered to each other. By contrast, in the genitive condition, the mean peak height of NP1 and NP2 stays at the same height and rises again on NP3. These results are depicted with an examplary pitch contour (cf. Figure 4.2) and with a pitch track averaged over all speakers (cf. Figure 4.3). Statistical analysis confirmed these observations. A multilevel model (Bates & Sarkar, 2013) was fitted using crossed random factors subject, item and condition (GEN, DAT) as fixed factors. The analysis relied on f0 maximum as a dependent variable. The statistical comparison revealed a significant effect of f0 maximum within interval-NP1 and interval-NP3 (cf. Table 4.4).

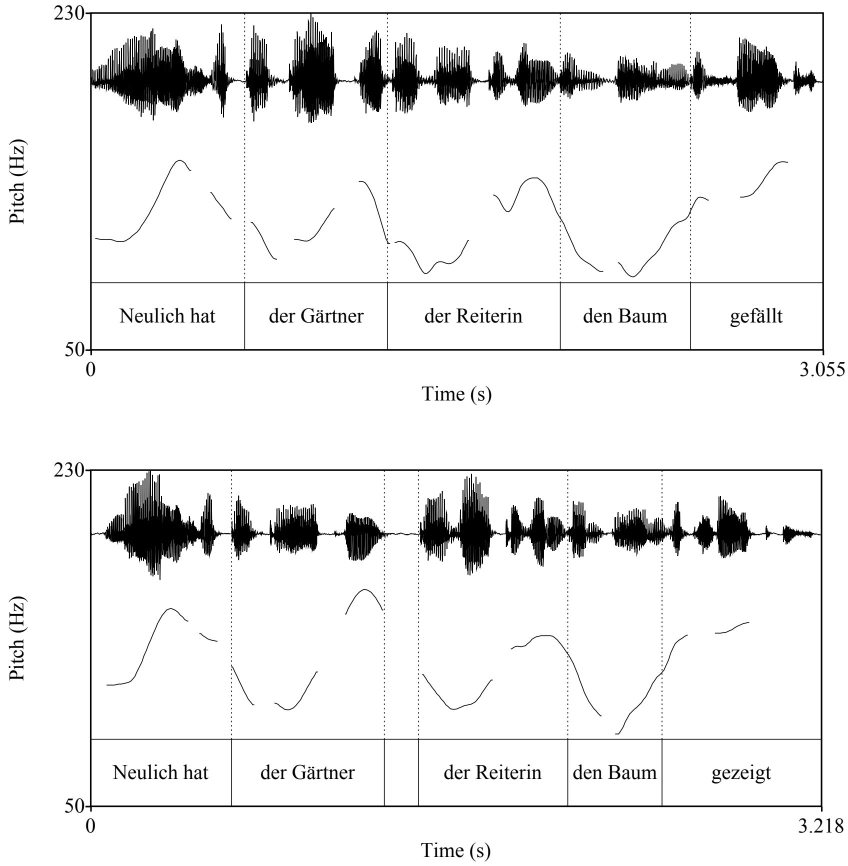

Figure 4.2: Example pitch contour of a genitive sentence (upper panel). Sentence *Neulich hat der Gärtner der Reiterin den Baum gefällt.* 'Recently the gardener of the horsewoman chopped a tree.' Example pitch contour of a dative sentence (lower panel). Sentence *Neulich hat der Gärtner der Reiterin den Baum gezeigt.* 'Recently the gardener showed the horsewoman a tree.'

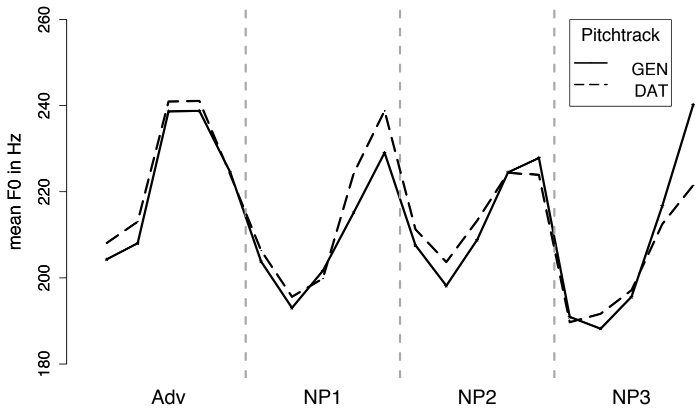

Figure 4.3: Mean pitch track, averaged over all speakers, of the ambiguous sequence, i.e., the adverbial phrase, the first NP, the second NP and the third NP for both, the genitive (solid line) and the dative condition (dashed line)

Constituent	Comparison	Means (f0 max (Hz))		t-scores
ADV	GEN vs DAT	227.16	228.94	<1
NP1	GEN vs DAT	237.39	250.81	3.09*
NP2	GEN vs DAT	248.31	246.67	-0.19
NP3	GEN vs DAT	255.12	237.67	6.92*

Table 4.4: Mean (f0max (Hz)) on the first four constituents of genitive and dative sentences as well as their statistical comparisons by means of t-tests.

To summarize, for the genitive condition, Figure 4.2 (upper panel) shows that the adverbial phrase, and all noun phrases are associated with a rising pitch accent (L*H). The right edge of the matrix clause is associated with a high phrase boundary tone (Hp) indicating a sentence continuation (Cruttenden, 1986; Féry, 1993). Similarly, for the dative condition, Figure 4.2 (lower panel) depicts that the adverbial phrase and the succession of all three noun phrases is associated with a rising pitch accent (L*H) as well. The following verb carries a high phrase boundary tone (Hp) as well signaling the continuation of the sentence. The results of the phonological analysis that correspond to both sentence conditions are summarized in Table 4.5 below and show that the numbers, distribution and type of the pitch accents are identical in both sentence conditions.

Experiment	Condition	NP1	NP2	NP3	VER
φ-boundary	GEN	L*H	L*H	L*H	Hp
	DAT	L*H	L*H	L*H	Hp

Table 4.5: Summary of the prosodic structure for both sentence conditions

Prosodic phrasing

Prosodic phrasing emerges from the syntax-prosody mapping, see section 2.1.3. This mapping results in prosodic domains that are, following (Féry, 2010b), called prosodic phrases (p-phrases) and labelled as φ-phrases throughout this work. Each φ-phrase is supposed to have a head, which is implemented as a pitch accent, and every pitch accent reflects the head of a φ-phrase. In the same way as syntax is unrestricted and does not constrain the (potentially) maximal given number of embedded sentences within one sentence, the number of prosodic domains within one sentence should be equally undefined. Therefore, prosodic phrasing is regarded as unconstrained and described as recursive in this thesis. Based on the results of the phonetic measurements, the phrasal properties are suggested below, separated for each condition.

Prosodic phrasing of the genitive condition: (NP1 (NP2)φ)φ

The prosodic contour indicates that the utterance in (15a), here repeated as (16a) consists of four φ-phrases that are recursive and themselves contain embedded φ-phrases. We assume that in such a structure the entire sentence represents one φ-phrase. This φ-phrase is not dominated by any higher φ-phrase, and, therefore is at the same time an ι-phrase. This ι-phrase is divided into four separate φ-phrases. The first φ-phrase carries one bitonal pitch accent L*H on *neulich*, 'recently'. The second φ-phrase includes NP1 and NP2, both of which carry a bitonal accents, the first one on *Gärtner*, 'gardener', the second one on *Reiterin*, 'horsewoman'. The third φ-phrase is the verbal phrase *den Baum gefällt*, 'the tree chopped'. This verbal phrase carries a further, embedded φ-phrase, which is the object of the verb *den Baum*, 'the tree'. This object φ-phrase carries the bitonal pitch accent L*H and is the head of its own phrase and, at the same time, the head of the φ-phrase that corresponds to the entire verbal phrase. The verb itself carries a high phrase boundary tone (Hp). The last φ-phrase is represented by the relative clause *der morsch war*, 'that decaying was', realized with a bitonal falling pitch accent H*L associated with the lexically stressed adjective. The last word of the utterance, *war*, 'was', carries the low boundary tone L% of the entire ι-phrase.

(16a)

L*H L*H L*H L*H Hp H*L L%

[[NEUlich hat]φ [der GÄRTner [der REIterin]φ]φ [[den BAUM]φ gefällt]φ [der MORSCH war]φ]φ=ι

Recently the gardner of the horsewoman chopped the tree, that was decaying.

Prosodic phrasing of the dative condition: (NP1)φ (NP2)φ

In contrast to the genitive condition which consists of four φ-phrases, the dative structure consists of five φ-phrases, which are again recursive and so contain further embedded φ-phrases, as schematically presented in (16b). The entire sentence is one φ-phrase of its own. As shown in (16a), this φ-phrase is not dominated by a higher φ-phrase and so corresponds with an ι-phrase. The entire ι-phrase is broken down into five single φ-phrases, each of which carrying a pitch accent. The first φ-phrase carries one bitonal pitch accent L*H on *neulich*, 'recently'. The second φ-phrase carries its bitonal pitch accent on *Gärtner*, 'the gardener'. The third φ-phrase carries the bitonal pitch accent on *Reiterin*, 'horsewoman'.

The fourth *φ*-phrase includes the entire verbal phrase *den Baum gezeigt*, 'the tree showed'. This verbal phrase carries a further, embedded *φ*-phrase, represented by the object *den Baum*, 'the tree'. This object *φ*-phrase carries the bitonal pitch accent L*H and is, as already shown in (16a), the head of its own phrase and, the head of the *φ*-phrase that corresponds to the entire verbal phrase. The verb *gezeigt*, 'showed' is associated with a high phrase boundary tone (Hp). The last *φ*-phrase is again represented by the relative clause *der morsch war*, 'that decaying was', carrying a bitonal falling pitch accent H*L on the adjective *morsch*, 'decaying', and a low boundary tone L% at the last word *war*, 'was' of the entire utterance.

(16b)

L*H L*H L*H L*H Hp H*L L%

[[NEUlich hat]*φ* [der GÄRTner]*φ* [der REIterin]*φ* [[den BAUM]*φ* gezeigt]*φ* [der MORSCH war]*φ*]*φ*=*ι*

Recently the gardner showed the horsewoman the tree, that was decaying.

Summary

Summarizing the results, we find that speakers successfully disambiguate the genitive and dative construction by prosodic means such as *prefinal lengthening, pause duration*, and a difference in scaling the high tones of the pitch accents. In particular, the results of the phonological analysis reveal that all NPs have a rising tonal pattern (L*H) showing a significant difference in pitch height on NP1 and NP3 between both sentence conditions. In addition, significant differences were found both in the duration of all NPs and in the pause duration between NP1 and NP2. These prosodic deflections indicate that speakers significantly altered their production and produced a different intonational phrasing signaling a difference in meaning, respectively.

It motivates a recursive prosodic phrase structure of the genitive condition appearing as follows: NP1 and NP2 are syntactically grouped together and form one single *φ*-phrase (without an interfering prosodic boundary), where NP2 is embedded within NP1. Within that phrase pitch stays at the same level and rises again on NP3. This declination reset indicates the beginning of a new phrase. Evidence for analyzing NP1 and NP2 as part of one single *φ*-phrase results from the durational analysis as well which revealed neither *prefinal lengthening* of NP1 nor a pause between NP1 and NP2 - phonetic correlates which are highly correlated with the occurrence of a prosodic phrase boundary.

In contrast, in the dative condition, NP1 and NP2 are not syntactically grouped together and are thus phrased separately (with an interfering boundary). The height of the f0 values on both NPs are successively lowered from one to the next phrase. This phenomena is referred to as downstep. Regular downstep is considered as the default pattern of high tone implementation in a 'neutral' sentence, where neutral is understood as a discourse-neutral realization, i.e. a broad focus reading. Downstep can be understood as a tonal effect taking place across *φ*-phrases and is inter alia a correlate of the number of prosodic phrases, as shown in Féry & Kügler (2008). The distribution of pitch accents in a downstep fashion is discussed in Liberman & Pierrehumbert (1984) for English, and in Féry (2010b) for German. Bruce (1977), Clements (1981) and Ladd (1990) suggest similar proposals for different languages as well. These phrasing differences are comparable to the ones reported in the literature on the phrasing of verbal arguments versus adjuncts in that verbal arguments are preferably indicated by boundaries, while embedded structures are phrased together and mainly occur without boundaries (Augurzky & Schlesewsky, 2010).

4.2 Perception experiments

The production experiment revealed prosodic differences for three prosodic correlates of phrasal boundaries: *prefinal lengthening, pause duration*, and *f0* for the two different meanings of the syntactically ambiguous sentence. The three prosodic cues are well documented in the literature in signaling prosodic phrase boundaries, and many studies have demonstrated that listeners use these kinds of prosodic differences to determine speakers' intended meaning for ambiguous utterances (Lehiste, 1973; Lehiste et al., 1976; Price et al., 1991). Many scholars even go one step further and argue that listeners not only use prosodic cues to commit oneself to one specific reading of a syntactically ambiguous sentence, but indicate listeners' competence to predict the amount of so far unheard prosodic material on the basis of local prosodic cues. Grosjean (1983) presented English subjects with sentences carrying optional prepositional phrases of varying lengths. Dependent on the length of the prepositional phrase, sentences differed in their prosodic characteristics. Grosjean (1983) showed that based on these different prosodic characteristics of a sentence fragment, listeners are able to anticipate the number of words to the end of a sentence - obviously a relevant clue for the computation of the syntactic structure, see also Grosjean & Hirt (1996), for a replication of this study in English and French. Likewise, Snedeker & Trueswell (2003) found that prosodic information in the speech input is likely to drive listeners towards one reading of a syntactically ambiguous sentence even before encountering the ambiguous sequence. According to this, it seems to be a fruitful approach to examine the impact of each individual prosodic correlate on sentence comprehension - as opposed to the influence of combined prosodic cues - by using the predictive role of prosody in a sentence completion experiment. This way, the as yet lacking information with respect to the influence of individual prosodic correlates on sentence disambiguation in German can be tackled.

In the following, six perception experiments will be reported that have two goals. First, we ask whether listeners use the prosodic differences -calculated in production- for their perception in order to disambiguate the sentences. Second, we ask whether *prefinal lengthening, pause duration* or the effects of *intonation* are used in resolving German ambiguities by neutralizing each of the three prosodic cues. The aim is to examine whether listeners' performance on sentence completion tasks remain the same or are degraded as a result of the loss of one or more of these prosodic parameters. In this way, information concerning the perceptual weighting of these cues can be elucidated.

4.3 Experiment 1: Sentence disambiguation using original stimuli

Based on prosodic disambiguation studies, we hypothesize for Experiment 1 that prosodic correlates of the speech signal will influence and guide listeners' disambiguation. An effect of prosody is thus anticipated: in a forced-choice experiment, the prosodic correlates associated with the genitive condition should influence listeners in such a way that they disambiguate the auditory sequence more often towards the genitive interpretation. In case of the dative condition, the corresponding prosodic correlates should lead participants more often to opt for the dative sentence continuation as opposed to the genitive one.

According to the parsing principle *Minimal Attachment* we expect to find an overall processing preference of the dative condition.

Material

20 sentences of each case condition were pronounced by a trained speaker incorporating the intonational patterns of the production study. A congruence in prosodic structure of these recordings with the production data were confirmed by phonetic measurements of f0, prefinal lengthening, and pause duration for the temporarily ambiguous region. Sentences that did not match the mean intonational contour of the production study were replaced by new recordings. In this way we guaranteed a consistent data set realized by one speaker. Subsequently, the disambiguating part of all target sentences (i.e. the verb and the following relative clause) was cut off. Each fragment was equated in loudness at the level of 70 db using a scale intensity praat script. The resulting sentence fragments were taken to test the extent to which listeners use the information in auditory parsed sentence fragments to predict upcoming entities.

Method

Each sentence fragment was presented to 20 listeners via headphones. None of the listeners participated in the previous production experiment. In a forced-choice task, participants listened to each sentence fragment up to and including NP3, and then were asked to complete the sequence by choosing one of the two continuations (either the genitive or dative verb) offered in the questionnaire. Listeners were aware that both choices were plausible but were forced to choose the most appropriate continuation depending on the way the heard sequence was uttered. There is an advantage of presenting possible sentence continuations that are attributable to either of the two conditions, as opposed to studies where listeners had to complete the auditory sequence with a completion that first came to their minds (Millotte et al., 2008). By presenting possible completions we avoided that participants gave sentence continuations that did not explicitly disambiguate the corresponding structures. The target sentences were intermixed with numerous filler sentences that varied in prosodic and syntactic structure, in order to discourage response strategies. Before the experiment began, participants performed a 3-items training phase.

Results

The results of the sentence completion experiment display a very clear pattern in that listeners recognize the different syntactic structures on the basis of prosodic cues (cf. Figure 4.4). Specifically, in 67% of the cases, the correct genitive verb was selected when a genitive sequence (i.e. genitive prosody) was presented. When a dative sequence (i.e. dative prosody) was presented, listeners selected in 93% the prosodically fitting dative verb. For the statistical, frequency-based analysis, we fit a multilevel model (**?**) using crossed random factors subject and item, and sentence sequence (GEN, DAT) as fixed factors. The analysis relied on the choice of verb completion as a dependent variable. The statistical comparison revealed a significant effect of the selected verb, in that, independent of the presented auditory sequence, listeners selected the prosodically correct sentence continuation significantly ($z=5.09$, $p>0.001$) more of-

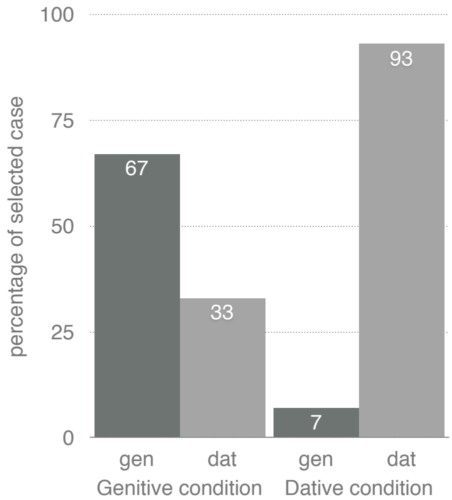

Figure 4.4: Percentage of the selected sentence continuation for the unmanipulated stimuli of Experiment 1.

ten, confirming the actual use of prosodic cues to identify the intended syntactic structure. Furthermore, the expected preference for the dative condition was verified.

4.4 Experiment 2: On the impact of duration on sentence disambiguation

Given that listeners reliably differentiated between the case ambiguous sentences by means of prosodic information (cf. Experiment 1), the following perception experiments address the issue of *which* prosodic cues listeners actually use when disambiguating the target sentences. In the current perception experiment, the controlled intonational manipulation of sentences' ambiguous constituents were investigated to examine whether the remaining prosodic information of the duration (prefinal lengthening and pause duration) was sufficient for listeners when disambiguating the sentences. In other words, the sentence fragments contained natural durational information and manipulated f0-information. Since the prosodic evaluation of the production data revealed no systematic difference of the tonal structure (besides a difference in pitch accent scaling among both sentence conditions), we hypothesize that the removal of the intonational contour would only slightly influence listeners' disambiguation ability, if at all. By removing the intonational contour, we can shed light on the exclusive impact of durational information on syntactic ambiguity resolution.

Material

The sentence material used in Experiment 1 was taken and *f0*-manipulated by removing all pitch cues using a praat script (Boersma & Weenink, 2013). To remove all pitch cues, the fundamental frequency of all voiced portions of the sentences was set to a constant value of 120 Hz, resulting in a flattened

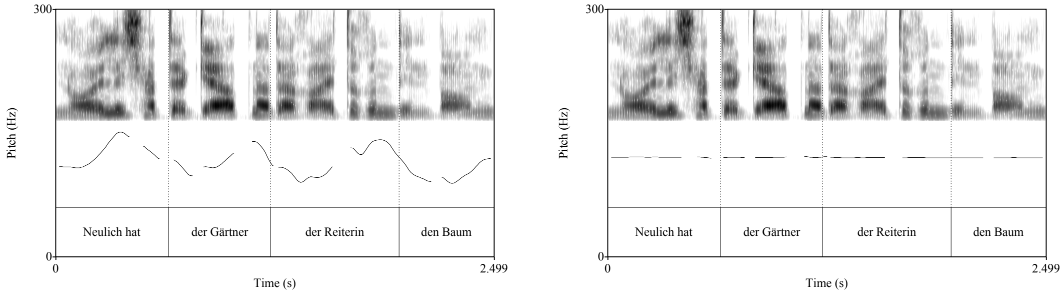

Figure 4.5: Illustration of an unmanipulated stimuli (left) and a f0-manipulated stimuli (right). The graphics picture the ambiguous sequence up to and including NP3.

f0-contour (cf. Figure 4.5, right graphic). At the same time, the durational properties of the sentences were not changed.

Method

The f0-manipulated target sequences were again intermixed with filler sentences from four unrelated experiments to avoid response strategies. 20 participants took part in the perception experiment, they were all native speakers of German, and reported no hearing impairment. They were asked to listen to each manipulated sequence via headphones and then select either one of the two verbs to complete the sentences in the forced-choice manner described above. They did not participate in any of the preceding experiments.

Results

On the basis of durational information, listeners selected the prosodically fitting sentence continuation in 65% of the genitive sequences. When confronted with dative sequences, listeners selected the correct verb in 87% of the cases (cf. Figure 4.6). The statistical, frequency-based analysis revealed a significant effect of the selected verb ($z=7.50$, $p<0.001$) between both sentence conditions. Comparing these results with the outcome of perception Experiment 1 (cf. Experiment 1, Figure 4.4), a similar pattern emerges, in that, independent of the presented auditory sequence, listeners selected more often the prosodically fitting verb. A statistical comparison of Experiment 1 with Experiment 2 revealed a non-significant verb selection for both conditions ($z=-0.984$), for the dative conditions ($z=-1.313$), as well as for the genitive conditions ($z=-0.109$), indicating that the prosodic manipulation performed on the sentence material in Experiment 2 did not affect the verb selection. Therefore, as a result of Experiment 2, we may reach the preliminary conclusion that the prosodic disambiguation observed in Experiments 1 and 2 reflects a process that does necessarily require the information of f0. In other words, the uninformative cues provided by f0 in Experiment 2 did not affect listeners disambiguation ability, confirming the hypothesis that the removal of the intonational contour does not affect listeners' disambiguation ability.

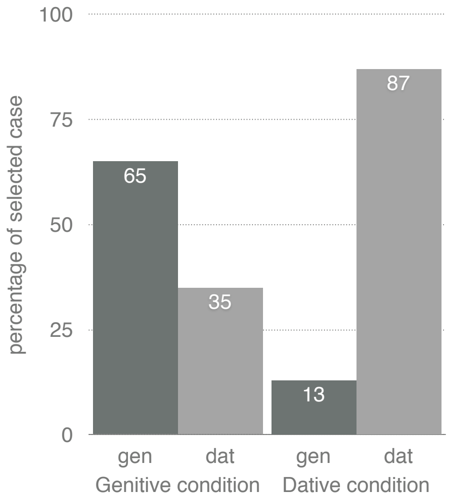

Figure 4.6: Percentage of the selected verb continuation for f0-manipulated sentences (i.e., sentence fragments contained natural segmental and pause duration and manipulated f0-information (120 Hz)) of Experiment 2.

4.5 Experiment 3: On the impact of f0 on sentence disambiguation

A third perception experiment will supplement the findings of Experiment 2 by scrutinizing the effect of f0. The central motivation of Experiment 3 is to further explore the so far missing contribution of f0 on ambiguity resolution. For the present experiment, the sentence material was manipulated so that it contained ambiguous durational information. As a result, the f0 would be the only decisive cue. Juxtaposing the results of Experiment 2 with the results of Experiment 3 makes the impact of f0 transparent. Based on the results of the production experiment, Experiment 1 and 2, the prediction for Experiment 3 is as follows: since listeners disambiguated the ambiguous sequences even in the absence of cues provided by f0 (Experiment 2), it is hypothesized that the meaning of the two sentence conditions cannot be ascertained with original f0 correlates and manipulated (ambiguous) durational information due to the small difference in f0 deflection, as evidenced by the production experiment above. To test that prediction, we used sentence material carrying original f0 information and manipulated (ambiguous) durational information.

Material

The sentence material used for Experiment 3 contained natural f0 cues and ambiguous durational information to rule out an influence of duration. Specifically, the durational cues were manipulated such that the word and pause duration of each ambiguous sentence pair carried a mean value. For that, the word and pause durations for each ambiguous sentence pair were measured; out of the duration data, a mean value for each word and pause (P1) per sentence pair was calculated. Then the stimuli were modified by stretching or compressing the word and pause durations in order to match the mean value (see Lehiste et al. (1976) for a description of this method). Modifications to the stimuli for each sentence pair were done using praat (Boersma & Weenink, 2013). Table 4.6 and 4.7 illustrate the results of the manipulation procedure and show the original duration and the calculated mean values collapsed across conditions for 2 of the 20 target sentence pairs (cf. sent. pair 1. and 2.). Table 4.6 (last column) presents the results for the word manipulation, Table 4.7 (last column) shows the pause manipulation. Appendix C contains a table showing the durational manipulation for the entire stimulus set used in Experiment 3. Note that a pause manipulation of the second pause (P2) was not performed since the mean pause values of P2 between both condition did not differ significantly.

sent pair	manipulated words	condition	original duration	changed to
1.	|NP1| |NP2| |NP3|	DAT	650 700 470	555 725 460
		GEN	460 750 450	460 750 450
2.	|NP1| |NP2| |NP3|	DAT	780 550 420	605 635 455
		GEN	430 720 490	605 635 455

Table 4.6: Word duration manipulation in ms of genitive and dative sentences

sent pair	manipulated pause	condition	original duration	changed to
1.	|P1|	DAT	0.13	0.075
		GEN	0.02	0.075
2.	|P1|	DAT	0.04	0.035
		GEN	0.03	0.035

Table 4.7: Pause duration manipulation in ms of genitive and dative sentences

The target sequences were interspersed with filler sequences and randomized in order to discourage response strategies. 20 participants took part in the perception experiment. None of the subjects participated in either of the preceding experiments. They were native speakers of German, reported no hearing deficits, and were reimbursed for participation with course credit points. Each participant was presented either one of the two sentence lists. The participants listened to each sequence via headphones and were asked to select one of the two offered verbs in order to complete the fragments.

Results

Figure 4.7 presents the results of Experiment 3 and shows that listeners were not able to disambiguate the auditory sequence on the basis of f0 cues. Instead, they resolved the ambiguity towards the dative structure, independently of the presented condition. This pattern emerged in 59.5% when listeners were confronted with genitive sequences and in 91% of cases when listeners were confronted with dative sequences. The frequency-based statistical analysis showed a significant main effect of CONDITION on the verb selection ($z=8.474$, $p>0.001$); independently of the presented auditory condition, listeners disambiguated the sequences towards the dative structure. In other words, if word and pause durations were erased in the sense that they did not provide decisive information with respect to either condition, the remaining f0 information by itself did not allow listeners to disambiguate the sequences. There is however, a considerably difference in the amount of given dative answers between both conditions, i.e. 59.5% as opposed to 91% which may be attributable to the difference in f0, but still, f0 on its own is not a sufficient

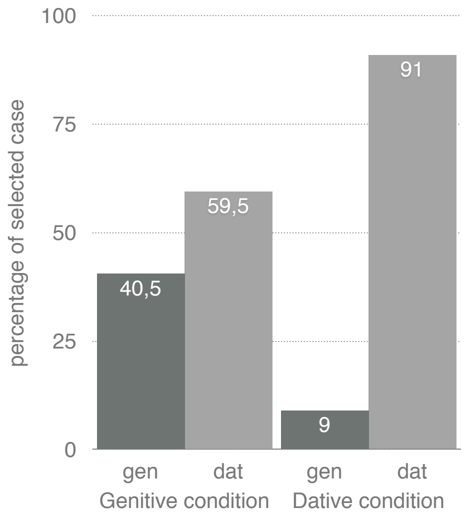

Figure 4.7: Percentage of the selected verb continuation for durational-manipulated sentences (i.e. sentence fragments contained natural f0 cues and manipulated word and pause duration) of Experiment 3.

prosodic correlate that lets listeners disambiguate the structures explicitly. This result is in accordance with the outcome of Experiment 2 that f0 did not substantially contribute to listeners' disambiguation. Furthermore, this result confirms the conclusion of Experiment 2, that it is the durational information that listeners rely on when committing to one of the two structures. This assumption will be further explored in Experiment 4, where the strength of the durational cue is scrutinized.

4.6 Experiment 4: On the strength of the duration cue for sentence disambiguation

As much as the above results support the assumption that word and/or pause duration is a necessary cue for listeners in assigning a syntactic structure, it is only preliminary evidence and crucially requires the examination of its effect on sentence disambiguation in more detail. By acknowledging an impact of duration, one might assume that it is possible to shift answering patterns in a forced-choice paradigm such that for the more preferred structure carrying the durational properties of the less preferred structure, listeners would compute more often the structure that is associated with the less preferred reading and therefore would frequently select verbs that would complement the less preferred reading. Specifically, we would expect for dative sequences carrying genitive-like durational properties, that listeners would compute the genitive structure and thus are more likely to opt for genitive than dative verbs to complete the auditory sequence. If that is indeed the case, independent evidence on the exclusive role of duration for disambiguation is provided.

Material

The genitive sequences were not changed, and carried original durational properties (see Experiment 1). Only dative sequences were manipulated individually. For each dative sequence, the word and pause durations were stretched or compressed to match the original word and pause properties of the genitive counterpart. This manipulation procedure resulted in a stimulus set carrying only genitive-like word and pause duration properties. The modifications to the dative stimuli were done using praat (Boersma & Weenink, 2013). Table 4.8 and 4.9 illustrate the results of the durational manipulation of the dative sequences for two sentence pairs referred to as sent pair 1. and 2. Table 4.8 presents the results for the word duration manipulation, Table 4.9 depicts the pause duration manipulation of the the pause between NP1 and NP2 (referred to as P1) for each ambiguous interval, respectively. Appendix C contains two tables showing the word and pause manipulation carried out for all dative sequences in Experiment 4.

sent pair	manipulated words	condition	original duration	changed to
1.	|NP1| |NP2| |NP3|	DAT	650 700 470	460 750 450
	|NP1| |NP2| |NP3|	GEN	460 750 450	460 750 450
2.	|NP1| |NP2| |NP3|	DAT	600 550 370	370 570 440
	|NP1| |NP2| |NP3|	GEN	370 570 440	370 570 440

Table 4.8: Word duration manipulation in ms of dative sentences

sent pair	manipulated pause	condition	original duration	changed to
1.	|P1|	DAT	0.13	0.02
	|P1|	GEN	0.02	0.02
2.	|P1|	DAT	0.04	0.03
	|P1|	GEN	0.03	0.03

Table 4.9: Pause duration manipulation in ms of dative sentences

Results

The results of the sentence completion experiment are presented in Figure 4.8 (right graphic). For the unmanipulated genitive sequences, listeners selected in 62% the correct genitive verb and in 38% the dative verb. For the durational manipulated dative sequences, the dative verb was selected in 60%, and the genitive verb in 40%. A statistical analysis revealed that the presented conditions (GEN, DAT) did not have a significant influence on the verb choice (z=-0.879).

Comparing the results with the outcome of Experiment 1 (cf. Figure 4.4 and the statistical analysis) a different pattern emerged in that the presented condition (GEN, DAT) had a significant influence on the correct sentence continuation. Here, the statistical analysis for Experiment 1 revealed a main effect of the sentence condition on the verb choice (z=5.09, p>0.001). Thus, the experimental manipulation in Experiment 4 changed the answering pattern in the forced choice task compared to the processing of the unmanipulated sentence material. Unexpectedly, in Experiment 4, there is no genitive preference

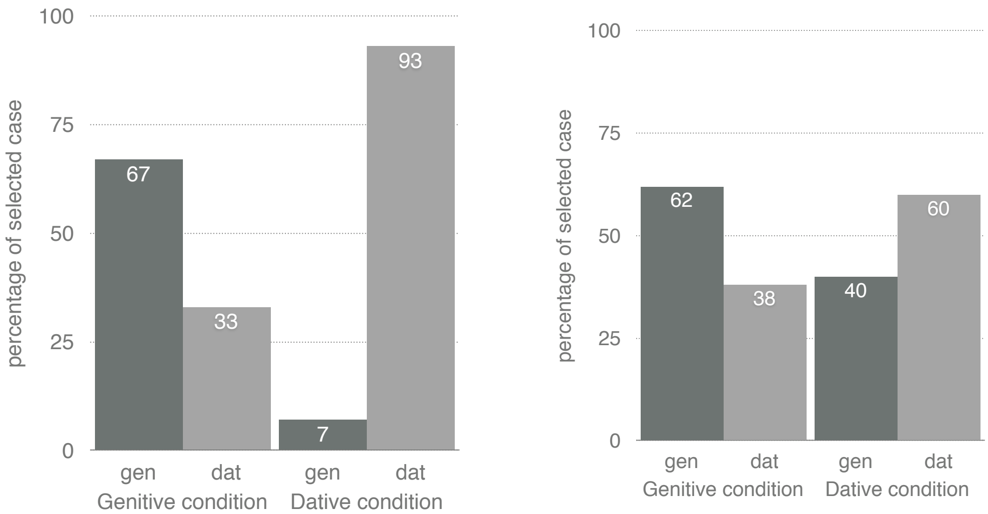

Figure 4.8: Percentage of the selected verb continuation in Experiment 1 (left graphic) compared to Experiment 4 (right graphic).

for a genitive-like manipulated dative sequence. However, it is the difference of the response pattern for the dative sequence in Experiment 1 compared to the current Experiment 4, that is crucial. A closer look at the result pattern for the dative conditions reveals the response change: the selection of the dative verbs in presented dative sequences dropped from 93% (Experiment 1) to 60% (Experiment 4). Similarly, the selection of genitive verbs in a presented dative sequence increased from 7% (Experiment 1) to 40% (Experiment 4) (cf. Figures 4.8). This difference was statistically confirmed with a post hoc test by fitting a multilevel model using SUBJECT and ITEM as random factors, and CONDITION as fixed factor. The analysis relied on the VERB as dependent variable. The statistical comparison of the verb selection for the dative condition between Experiment 1 and 4 revealed a significant interaction of the experimental manipulation ($z=-573$, $p>0.001$). The findings clearly corroborate the overall preference of the dative structure which cannot be inverted but interfered by changing the durational properties of a dative sequence into a genitive-like sequence. The prosodic cue *duration*, here meant as word and pause duration in one, appears to be a cue that listeners rely on when committing to one of two ambiguous sentence structures.

4.7 Experiment 5: On the influence of prefinal lengthening

In order to scrutinize the effects of *prefinal lengthening* from *pause duration* and investigate their independent influence on sentence disambiguation, two last perception experiments were conducted to ascertain whether the two prosodic correlates influence listeners' disambiguation ability in similar ways. Juxtaposing the results of the two experiments that examine (i) the impact of *prefinal lengthening* and (ii) the influence of the pause duration for disambiguation, makes the individual influence of each correlate transparent and allows us to gauge their impact on sentence parsing. To test the influence of *prefinal* *lengthening*, we left the word duration in both conditions intact and only manipulated the pause duration within the dative condition.

Material

Just as in Experiment 4, the genitive sequences remained unchanged. For the present Experiment 5, only pause properties of the dative sequences were manipulated individually, segmental duration properties were not changed. We modified the original dative stimuli used in Experiment 1 by changing the pause values of P1, i.e. the pause duration between NP1 and NP2, of each dative sentence towards the pause properties of the genitive sentences, so that the pause durations in both conditions were identical. The pause values of P2, i.e. the pause duration between NP2 and NP3 remained unchanged as their mean values across all sentences did not reveal significant differences. Table 4.10 shows the pause manipulations for two ambiguous sentence pairs. Appendix C contains a table showing the pause manipulation carried out for all dative sequences in Experiment 5. We chose to do this modification, rather than stretching the duration of the genitive pause in order to match them with the dative pause to counterbalance the preference for the dative structure, see also footnote 1 on a small discussion on that. The modifications to the stimuli of the dative condition were done using praat (Boersma & Weenink, 2013).

sent pair	manipulated pause	condition	original duration	changed to
1.	|P1|	DAT	0.13	0.02
	|P1|	GEN	0.02	0.02
2.	|P1|	DAT	0.04	0.03
	|P1|	GEN	0.03	0.03

Table 4.10: Pause duration manipulation in ms of dative sentences

Results

The results of the pause manipulation are shown in Figure 4.9 (right graphic). Listeners are able to successfully disambiguate the genitive as well as the dative sequences. Specifically, in 59.75% the correct genitive verb was selected when a genitive sequence was presented. When a dative sequence was presented, listeners selected the prosodically fitting dative verb in 80.5%.

The statistical, frequency-based analysis, revealed a significant effect of the selected verb, independently of the presented auditory condition, that is, listeners selected the prosodically correct sentence continuation significantly ($z=5.73$, $p>0.001$) more often. The response pattern of Experiment 5 is very much in line with the findings of Experiment 1, where listeners were able to disambiguate both sentence conditions, albeit on a higher level for the dative structure (see Figure 4.9, left graphic). The occurrence of a disambiguating effect in Experiment 5 leads one to suspect that the results observed for the verb selection are due to the word duration inherent in both sentence conditions. A disambiguating effect solely due to f0 can be excluded (see Experiment 2 and 3), likewise a disambiguating effect based on pause duration can be factored out due to the ambiguous pause information within both conditions (see pause manipulation of the current Experiment 5). To summarize, the results of the current sentence completion

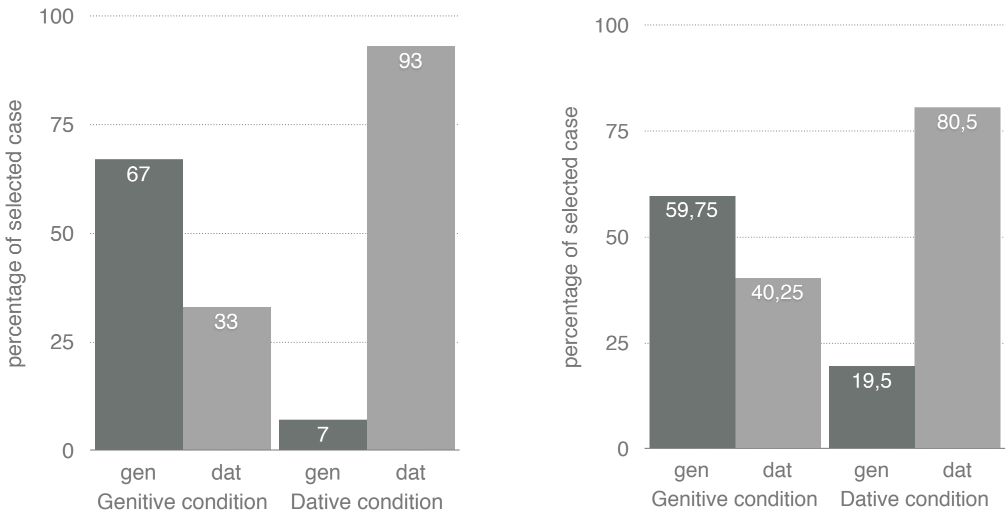

Figure 4.9: Percentage of the selected verb continuation in Experiment 1 (left graphic) compared to Experiment 5 (right graphic). In Experiment 5 pause-manipulated sentences were tested (i.e. sentence fragments contained natural word duration and manipulated pause duration).

experiment give evidence that listeners use *prefinal lengthening* to discriminate between the genitive and dative condition.

4.8 Experiment 6: On the failure of pause duration on disambiguation

Based on the previous perception results revealing a general influence of *duration* on listeners' disambiguation (cf. Experiment 2), and in particular showing the successful disambiguation based on *prefinal lengthening* only (Experiment 5), we hypothesize that the information of *pause duration* is not the most essential prosodic correlate listeners make recourse to when choosing among the competing syntactic realizations for an intended meaning. Therefore it is predicted for Experiment 6 that listeners' disambiguation ability decreases when *pause duration* is the only decisive correlate listeners can draw on during sentence disambiguation. To test the individual influence of *pause duration*, a perception study using manipulated word duration was carried out. This way, we disentangle whether *pause duration* and *prefinal lengthening* contribute to German ambiguity resolution differently.

Material

The original sentence material used in Experiment 1 was manipulated in two steps. First, for each ambiguous sentence pair, the word duration was measured; out of the resulting values the mean duration for each word of the ambiguous region was calculated. Then, the word duration was adjusted according to the mean duration of every constituent in the sentence pairs (NP1, NP2 and NP3). As a result, each manipulated sequence carried the mean of the length of the original sounds. Furthermore, to avoid response strategies favouring the dative interpretation, the pause duration (P1) of the dative condition was manipulated in a second step as follows: the pause duration values of the dative condition were reduced to match the pauses in the genitive condition.[1] Table 4.11 shows the original word duration and the mean values collapsed across conditions for two sentence pairs (last column). Table 4.12 depicts the pause manipulation performed on the dative conditions, and Appendix C contains two tables showing the durational manipulation for the entire stimulus set used in Experiment 6.

The manipulated sequences were interspersed with filler sequences and were auditorily presented via headphones to 20 participants. Similar to the preceding perception studies, participants did not take part in either of the previous experiments, they furthermore reported no hearing deficits and received course credit point for participation. After a 3-item-training period they were asked to listen to each manipulated sequence and then complete the fragments by choosing one of two offered verbs.

manipulated words	condition	original duration	changed to
|NP1| |NP2| |NP3|	DAT	650 700 470	555 725 460
	GEN	460 750 450	555 725 460
|NP1| |NP2| |NP3|	DAT	600 550 370	485 560 405
	GEN	370 570 440	485 560 405

Table 4.11: Segmental duration manipulation in ms of genitive and dative sentences.

sent pair	manipulated pause	condition	original duration	changed to
1.	|P1|	DAT	0.13	0.02
	|P1|	GEN	0.02	0.02
2.	|P1|	DAT	0.04	0.03
	|P1|	GEN	0.03	0.03

Table 4.12: Pause duration manipulation in ms of dative sentences.

Results

The results of the durational manipulation are depicted in Figure 4.10 on the right. It is evident, that independent of the presented auditory sequence, listeners choose the dative continuation significantly more often than the genitive verb continuation (selection of a dative sentence continuation: 83% for dative sentences and 85% for genitive sentences) and thus fail to disambiguate the sentences. A statistical comparison was carried out by fitting a multilevel model using SUBJECT and ITEM as crossed random factors, and CONDITION as fixed factor. The analysis relied on the VERB as dependent variable and revealed a non-significant verb selection independent of the sentence condition ($z=1.61$, $p=0.039$). This result indicates three things, firstly, when word duration is averaged across conditions, listeners fail to compute the correct sentence structure. Secondly, information on the pause duration (which was the

[1] The alternative method of keeping the original pause durations for P1 in both conditions could have been taken as immediate indication to interpret the fragments per se as dative constructions. The second pause duration P2 was not manipulated as the conditions did not significantly differ in their P2 values.

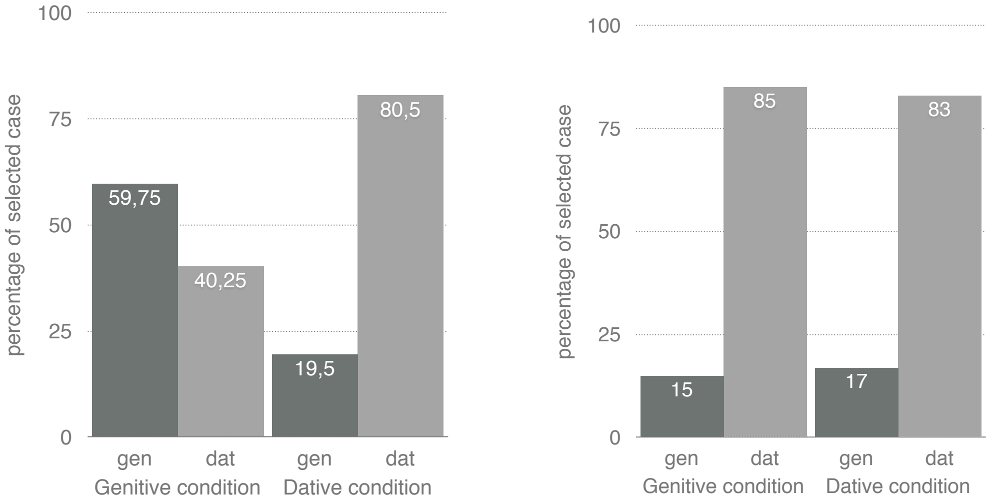

Figure 4.10: Percentage of the selected verb continuation in Experiment 5 (left graphic) compared to Experiment 6 (right graphic).

pause duration of the genitive condition throughout all stimuli) prevented listeners from correctly parsing the sentences. And thirdly, isolated pitch information (which was the only unambiguous prosodic cue) did not guide listeners towards the correct sentence interpretation. These results differ clearly from those found in Experiment 5, mainly in the distribution of given answers for the genitive condition. While in Experiment 5, in 59.75% of the cases, the genitive verb was selected, it was only selected in 15% of the cases for the same condition in Experiment 6. Note that the only prosodic difference between Experiment 5 and 6 was the information of the word duration, that was available and meaningful in Experiment 5 but not in Experiment 6. The results coming from Experiment 5 and from Experiment 6 clearly show a strong influence of *prefinal lengthening* for sentence disambiguation. While a *mean* value of the word duration of all constituents in both conditions prevent listeners from disambiguating the two alternative syntactic representations (Experiment 6, Figure 4.10 right graphic), *original* word durations seem to be a crucial cue that listeners effectively use for disambiguation (Experiment 5, Figure 4.10 left graphic). Additionally, the comparison of the two sets of results show that different response patterns are obtained depending on the prosodic manipulation. Although listeners had to accomplish the same task, the different responses can be attributed to the difference in the material used.

4.9 Summary of the experimental evidence

The production and sentence completion experiments examined the processing of case ambiguous sentences in German in view of the hypothesis that these constructions should be suitable for eliciting a cue weighting associated with boundaries of prosodic φ-phrases. From the nature of the constructions used to this end, essentially two interesting questions were addressed: namely, (i) which are the prosodic correlates speakers use to resolve the syntactic ambiguity under scruitiny, and (ii) are these correlates different from the ones listeners use when disambiguating the structures.

With respect to (i), we found that speakers significantly altered their production of the sentences in ways that were consistent with the intended structure by making use of changes in the height of the pitch accents, of prefinal lengthening and pause durations between both sentence conditions. With respect to (ii), the findings of six perception experiments revealed that listeners do not rely on all three prosodic correlates that speakers use for disambiguation, but rather make recourse to *prefinal lengthening* in resolving the ambiguity. Specifically, in Experiment 1 in which all prosodic cues were provided in the auditory signal, listeners were well able to disambiguate the syntactically ambiguous sequences, as was shown by the experimental findings of a more frequent genitive verb selection for the genitive sequences, and a more frequent dative verb selection for the dative sequences, respectively.

The first perception experiment using manipulated speech, reported here as Experiment 2, gives evidence that listeners were still able to disambiguate the sentences, even in the absence of cues provided by f0. For presented genitive sequences, more genitive than dative verbs were selected. Likewise, for the presentation of a dative sequence, more dative than genitive verbs were chosen by listeners. This indicates that the remaining prosodic information of the ambiguous auditory sequence, i.e. the collocation of word and pause duration, seems to be sufficient in order to successfully disambiguate the sequences towards a genitive or a dative reading, with a similar dynamic, as in Experiment 1. It is therefore apparent that, when considered together, Experiment 1 and 2 provide convincing evidence that the disambiguating effect in Experiment 1 is primarily based on the different durational properties of the two auditory sequences. It appears unlikely that the disambiguating effect in Experiment 1 is amenable to an explanation in which all prosodic cues are equally necessary, and are weighted as equally important by listeners to resolve the ambiguity. Moreover, the fact that the overall level of verb selection within both conditions was substantial and that it does not seem to be strongly related to f0, warrants two conclusions. First, the listeners do not seem to be any less capable of assigning the correct syntactic structure without f0 information and thus, do not heavily rely on that specific cue when parsing the sentences. Secondly, the method employed here is equally suitable for the analysis of *manipulated* sentences and allows the examination of a prosodic cue weighting.

The conclusions drawn from Experiment 2 were confirmed by Experiment 3, which examined the ambiguous sentences used in Experiment 1 by manipulating the word and pause duration simultaneously, so that they would not give decisive information to either of the structures. The results of the forced-choice experiment clearly show that the remaining unambiguous f0 information does not permit listeners to disambiguate the structures. Rather, listeners selected the dative sentence continuation more frequently and independently of the presented auditory sequence, indicating that the exclusive f0 cue is insufficient in order to disambiguate the sentences. This finding nicely parallels the outcome of Experiment 2 in providing convincing evidence that the disambiguating effect in Experiment 1 and 2 is not a result of using (or accessing) f0 information, but rather that of duration. The prosodic cue duration (i.e. word and pause duration collapsed) crucially influences the way in which the ambiguous constructions are processed in the disambiguating region.

Experiment 4 provided further perceptual evidence of the influence of duration on sentence disambiguation. Here, the dative sequences were manipulated so that each dative sequence carried word as well as pause duration properties of their genitive counterpart to examine to which degree duration can influence the response pattern of a dative sequence towards a genitive sequence. The verb selections in Experiment 4 reveal an ultimate preference for the dative structure, since genitive-like word and pause duration properties of the dative structure cannot lead to the less preferred alternative of the two possible syntactic constructions. In other words, the dative condition cannot be fully brought to elicit a genitive reading, even when the durational properties of a genitive structure are mapped onto the dative structure. However, the number of genitive verb selections in the context of a genitive-like durational manipulated dative structure significantly increased compared to the case of processing unmanipulated dative sentences. In this way, Experiment 4 demonstrated that word and pause duration influences the verb selection in the intended direction.

The combined findings of Experiment 5 and 6 provided perceptual evidence for the independence of word duration, i.e. *prefinal lengthening* from pause duration. In Experiment 5, the dative sentences were manipulated in such a way that each dative sequence carried pause duration properties of the genitive counterpart while leaving the word duration original. This manipulation procedure allowed us to test the exclusive impact of word duration on disambiguation. The results of the sentence completion experiment showed an explicit effect of disambiguation, i.e. for genitive sequences presented in the questionnaire, a higher selection of genitive than dative verbs were selected. Likewise, more dative than genitive verbs were chosen by listeners for the presentation of a dative sequence. In both conditions, correct sentence continuations were selected on a level similar to the one in Experiment 1, signaling that the prosodic cue *prefinal lengthening* is successfully used by listeners for disambiguation. When the information of word duration is neutralized and only pause information in both conditions is provided as a decisive cue (Experiment 6), listeners fail to disambiguate the sequences. These results speak in favor of a cue weighting, in which the phonetic cue most strongly associated with listeners perception of a φ-phrase boundary was *prefinal lengthening*.

4.10 General discussion

The finding that word duration appears to be a more crucial cue for disambiguation than pause duration, reinforces Goldman-Eisler's observation. She found that in English a pause is often used to signal cognitive activity rather than a syntactic boundary (Goldman-Eisler, 1972). Therefore, as a correlate of syntactic boundaries, pause duration appears to be less reliable than word duration. Another possible explanation for the word duration to be a decisive prosodic cue rather than pause duration is discussed in Schmitz (2009) who revealed that it is not the physical presence of a pause in the speech signal that triggers the perception of phrases and clauses. Rather, pauses seem to have a more complementary status: if other cues indicate a prosodic boundary, a pause is assumed to be present as well. Schmitz' study provided further evidence that the pause does not seem to be a necessary cue to identify φ- and ι-phrase boundaries in the speech stream, and argued that syntactic and/or prosodic units are often but not always demarcated by a pause. Furthermore, pauses may occur in other positions in the speech stream and also have different functions.

Recall that models of German intonation mainly differ with respect to representational and tonal aspects, see Kügler (2007). One of these differences refer to the controversial assumptions of the phrase accent which is postulated for German in the GToBI model based on Grice & Baumann (2002), but questioned by Féry (1993), who argued that the trailing tone of the last pitch accents of an intermediate phrase describes the tonal movement towards the intermediate phrase boundary. These are theoretical claims made in the literature which are based on the composition and ingredients of intonation. But none of the studies proved their claims perceptually. This work provides perceptual evidence that the identification of a φ-phrase boundary may be obscured or enhanced by the prosodic informativeness of the prosodic cue *prefinal lengthening* and implies a diminished or rather lacking tonal influence for its identification in perception. This observation speaks in favor of models in which φ-phrases are composed of pitch accents which are not obligatory followed by a phrasal boundary tone, implicating no need for an intermediate phrase boundary tone, as was already suggested by Féry (1993).

To summarize, the present findings shed new light on prosodic aspects of sentence comprehension. A production experiment and a sequence of six perception experiments examined the question of how speakers mark φ-boundaries, and which of the parameters they actually use, cue listeners' perception to disambiguate the ambiguous sequences. The experimental findings show evidence that the phonetic cues exploited by speakers in production to disambiguate the ambiguous sentences stand in contrast to the prosodic cues used by listeners to disambiguate sentences. While speakers altered the pitch peak height and used segmental and pause duration in order to disambiguate the structures, listeners seemed to rely primarily on *prefinal lengthening* when committing to one of the two ambiguous constructions. Thus, not all measurable differences in production are used in parsing prosody.

The dichotomy of phonetic cues exploited by speakers and those perceptual correlates used by listeners was already taken up by Heldner (2001) who emphasized that these two concepts do not necessarily relate to exactly the same physical reality, and agued "that listeners certainly do not have to use all the acoustic information available" (Heldner, 2001). He defined a perceptual correlate "as an acoustic feature either influencing a categorical distinction or affecting the perceived naturalness within a prosodic category."

In view of this definition, the present experiments provide strong support for a prosodic cue weighting on φ-boundaries, with *prefinal lengthening* being the most important correlate and *f0* being of minor importance for the perception of the φ-boundary, as only the difference of *prefinal lengthening* led to a categorical distinction between both sentence conditions.

It however, remains a possibility for the type of ambiguity involved in these sentences that at a different phrasing level (higher prosodic level), f0 and / or a combination of f0 and duration may be the prosodic cues listeners use to assign meanings. This possibility is examined in the following chapter of this thesis.

5

Prosodic cue weighting of I-Phrases

The level of the ι-phrase and its corresponding prosodic cue weighting is object of investigation in chapter 5. A production experiment using speech material containing an ι-phrase boundary was carried out to identify the prosodic correlates associated with the respective boundary, and revealed that similar to φ-phrase boundaries, *prefinal lengthening* and *pauses* appeared to correlate with an ι-boundary. In addition, the production experiment showed that ι-phrases are demarcated with a consistent boundary tone. A series of 5 perception experiments revealed that, from all measured boundary correlates, the boundary tone itself was identified as the most important prosodic cue to reliably distinguish the two syntactic structures of the syntactic ambiguity.

In the following, the main intonational features of an ι-phrase, as reported in the literature, are presented, before turning to the production experiment involving sentence material containing an ι-phrase boundary.

Generally, an ι-phrase more or less corresponds to a sentence and is appareled with a certain, grammatically defined, tonal inventory (Féry, 1993; Truckenbrodt, 2005). Every ι-phrase carries at least one pitch accent (Pierrehumbert, 1980; Féry, 1993). Usually, pitch accents within an ι-phrase undergo downstep, i.e. a lowering of each H peak relative to the preceding H peak, which has also been observed from Liberman & Pierrehumbert (1984); Prieto et al. (1996); Grabe (1998); Truckenbrodt (2004); Féry & Kügler (2008); Grice et al. (2009). Occasionally, the last nuclear (and commonly the strongest) accent within an ι-phrase may be reset to the tonal height comparable to the phrase-initial height. Resets are perceived as discontinuities and are exploited as a cue for strong boundaries (de Pijper & Sanderman, 1994). Phrase-initial resets count as an additional correlate of ι-phrases in certain southern dialects of German, and are used in very much the same way as boundary tones to signal an ι-phrase boundary Truckenbrodt (2002).

Futhermore, syllables preceding an ι-phrase boundary are lengthened as compared to syllables in non-phrase final positions. *Prefinal lengthening* is often considered to be universal, though the extent of lengthening varies among languages (Delattre (1966); Oller (1973); Hoequist (1983). *Prefinal lengthening* has been observed in English (Oller, 1973; Lehiste, 1973; Lehiste et al., 1976; Klatt, 1975; Ladd & Campbell, 1991; Wightman et al., 1992; Ferreira, 1993; Turk & Shattuck-Hufnagel, 2007), French (Delattre, 1966; Fletcher, 1991; Di Cristo & Hirst, 1997; Hirst, 1999; Tabain, 2003), Italian (Vayra & Fowler, 1992; D'Imperio & Gili-Favela, 2004; Avesani & Vayra, 2005), and German (Delattre, 1966; Kohler, 1983; Féry, 1993; Kuzla et al., 2007), to name a few languages.

At the end of an ι-phrase, tonal correlates of finality are realized, such as a falling intonation for declarative sentences conveying *finality*, or a rising intonation in questions or continuative utterances conveying *continuation*, implemented as a high (H%) or a low (L%) boundary tone, representing a difference in sentence mode (Schneider, 2012).

Additionally, ι-phrase boundaries can be followed or preceded by a pause. The correlation between syntax and pausing has been investigated by numerous researchers who observed that pauses mainly occur at major syntactic boundaries (Oller, 1973; Klatt, 1975; Cooper & Paccia-Cooper, 1980; Butcher, 1981; Ferreira, 1991; Horne et al., 1995; Yang, 2007). In most of these studies, pauses also tend to coincide with boundaries of higher-level prosodic constituents like ι-phrases or tone groups (Butcher, 1981; Gee & Grosjean, 1983; Ferreira, 1993; Krivokapic, 2007). There seems to be a positive correlation between pause duration and prosodic boundary strength (Strangert, 1991; Zellner, 1994; Choi, 2003).

There are other cues related to the voice quality that correlate with prosodic boundaries. A common observation is that the voice quality changes, in most of the cases, deteriorates at the end of a prosodic domain. Creaky voice is a frequent cue to end a prosodic phrase in English and many other languages, as already shown in Lehiste (1973).

Regarding syntactic factors, it has been pointed out that while syntactic structure does not fully determine prosodic structure, it does influence it to a large extent. Accordingly, there is ample evidence for cases where ι-phrases refer to syntactic clauses or interruptions of the main sentence (Ferreira, 1988; Gee & Grosjean, 1983; Selkirk, 2000). Furthermore, certain syntactic structures have been shown to form their own ι-phrases, for example, tag questions and unrestrictive relative clauses (Nespor & Vogel, 1986), or insertions in parenthesis (Dehé, 2014), as well as appositions.

Taken together, the intonational features of an ι-phrase suggest that there is little disagreement about the prosodic renderings. Furthermore they appear to be very similar to the prosodic correlates associated with φ-phrases. However, there is no evidence to suggest *which* prosodic information is used by listeners when processing ambiguous structures containing an ι-phrase boundary. Therefore, in the following, a line of perception experiments examined the prosodic cues used by listeners while processing ι-phrase boundaries. The perception results are required to (i) compare whether these cues differ from the ones that have been identified as crucial in processing smaller φ-phrase boundaries, and (ii) to eventually allow us to formulate a prosodic cue weighting on ι-phrase boundaries. The perception experiments are preceded by a production experiment aiming to identify the prosodic correlates associated with the sentence material under scrutiny.

5.1 Production experiment

A speech production experiment was set up to identify the prosodic cues used by speakers to demarcate ι-phrase boundaries in sentences like (17).

Material

16 sentences like (17) were devised that contained an adverbial phrase and a verb followed by two NPs. For the purpose of the production experiment, the contextual environment was systematically varied before the adverbial phrase and after NP2, such that NP2 could either be interpreted as a possessive

modifier of NP1 (17a), or as the object of the subsequent clause (17b), resulting in the fact that the ambiguous NP2 is part of one sentence in (17a), while it belongs to the second sentence in (17b). Each sentence pair used in the production experiment, including its corresponding context, is listed in Appendix A.

(17a) genitive-reading
Ständig isst der Kunde der Händlerin im Obstladen.
constantly eat the customerNOM the merchantGEN in the fruit shop
The customer of the merchant eats constantly in the fruit shop.

(17b) dative-reading
Ständig isst der Kunde. Der Händlerin vergeht das Lachen.
constantly eat the customerNOM. The merchantDAT does not like it
The customer eats constantly. The merchant does not like it.

The experimental sentences were controlled for number of syllables, stress pattern and gender. All verbs were monosyllabic, all NP1s were masculine disyllabic trochees and all NP2s were feminine trisyllabic trochees. The experimental sentences are highly sonorant to allow for a maximally accurate pitch analysis. The sentence material was interspersed with filler sentences from two unrelated experiments and fed into a linger presentation (tedlab.mit.edu/dr/Linger/). The sentences were pseudo-randomized for each subject, using the linger randomization settings such that sentences of the same condition did not appear adjacently and corresponding sentences of the two conditions had a maximal distance.

Participants

16 female native speakers of German participated in the experiment. All were undergraduate students at the University in Potsdam and were residents of the surrounding areas. They reported no speech or hearing impairment and received either course credits or were paid for participation. The speakers did not participate in any of the previous presented experiments.

Method

Each target sentence was presented together with its corresponding context visually on screen. The items were presented on a 15 zoll computer screen. Participants were asked to familiarize themselves with the sentences before reading them out aloud. Before the recording started, subjects performed three practice trials. In case of hesitations or false starts, participants were asked to repeat the sentence. Recordings took place in a sound-proof chamber equipped with an AT4033a audio-technica studio microphone, using a C-Media Wave soundcard at a sampling rate of 44.1 kHz with 16 bit resolution. Presentation flow was controlled by the experimenter, and participants were allowed to take a break whenever they wanted. 8 sentences were excluded due to hesitations, slips of the tongue or creakiness, so that a total of 504 sentences were included in the analysis. The target sentences were hand-annotated at the level of the constituent (see Table 5.1) and subjected to phonetic analysis using Praat software (Boersma & Weenink, 2013). The duration of the adverbial phrase, the verb, each NP and the pause between NP1 and NP2 were measured. Pitch analysis was conducted using a Hanning window of 0.4 seconds length with a default 10 ms analysis frame. The pitch contour was smoothed by using the smoothing algorithm (frequency band 10 Hz) to diminish microprosodic perturbations (Boersma & Weenink, 2013). An averaged pitch contour was created by dividing each constituent (see Table 5.1) into five equal-sized intervals and interpolating the mean F0 (in Hz) for each of these intervals over all speakers.

ADV	VERB	NP1	P	NP2
|Ständig|	|isst |	|der Kunde|	#	|der Händlerin|
constantly	eat	the customer	#	the merchant

Table 5.1: Illustration of the labeling positions, i.e the adverbial phrase (ADV), the verb (VERB), the first NP (NP1), the pause between NP1 and NP2 (P), and the second NP (NP2).

Results

A phonetic examination of the production data is carried out to identify the prosodic cues exploited by speakers. Word duration and f0 (maximum) values on the adverbial phrase, the verb, and both NPs were measured. Additionally, the pause duration between NP1 and NP2 in the genitive condition was calculated and compared with realizations of the dative reading. The following section presents the results of the phonetic evaluation. First, the prosodic contour is presented. Second, the results of the word and pause duration are reported. The results of the phonetic evaluation provide evidence for the phrasal structure of each condition, which is presented in the last section.

The prosodic contour

The following section presents the results of the analysis of the prosodic contour for each condition. First, the type and location of pitch accents for each defined interval in both conditions are illustrated with an examplary pitch contour (cf. Figure 5.1), and at a mean pitch track, averaged over all speakers (cf. Figure 5.2). Second, the difference in pitch accent scaling is indicated, and third, a summary of the phrasal structure for both conditions is offered.

Analogous to the data set tested in the previous chapter, a mean pitch contour, averaged over all 16 speakers was generated using a praat script. The procedure was identical to the one reported in chapter 4, section 4.1.4. The resultant averaged pitch contour is depicted in Figure 5.2. As for the genitive condition (solid line), there is a clear rising pitch contour on both NPs, indicating the realization of a rising pitch accent (L*H) on both NPs. In contrast, the dative condition (dashed line), shows a falling pitch accent (H*L) on NP1 and a rising pitch accent (L*H) on NP2. Two tonal differences are apparent: a rising versus falling accent on NP1, and a difference of the pitch accents height on NP2. Statistical analysis confirmed these observations. We fit a multilevel model (Bates & Sarkar, 2013) using crossed random factors subject, item and condition (GEN, DAT) as fixed factors. The analysis relied on f0 maximum as a dependent variable. The statistical comparison revealed a significant effect of f0 maximum within interval-NP1 (t-value = -14.45), and within interval-NP2 (t-value =13.11), cf. Table 5.2.

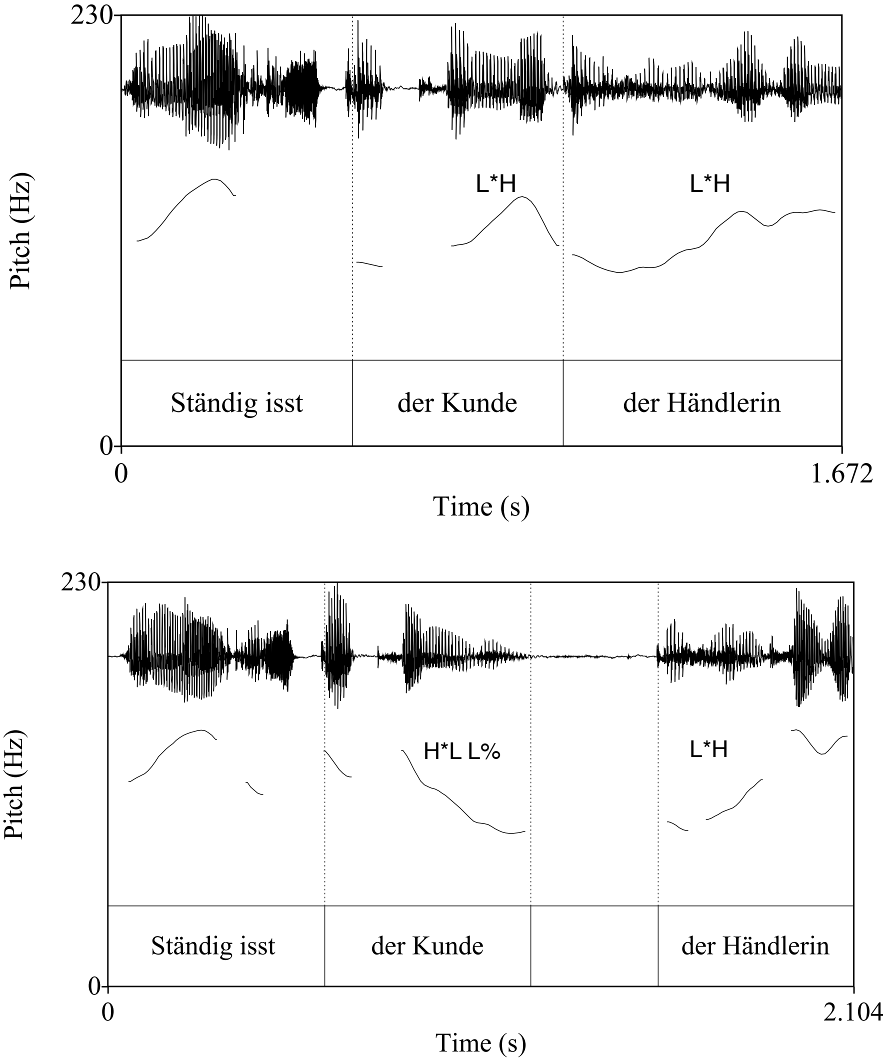

Figure 5.1: Example pitch contour of a genitive sentence (upper panel). Sentence *Ständig isst der Kunde der Händlerin (im Obstladen)*. 'The customer of the merchant eats constantly in the fruit shop.' Example pitch contour of a dative sentence (lower panel). Sentence *Ständig isst der Kunde. Der Händlerin (vergeht das Lachen)*. 'The customer eats constantly. The merchant does not like it.'

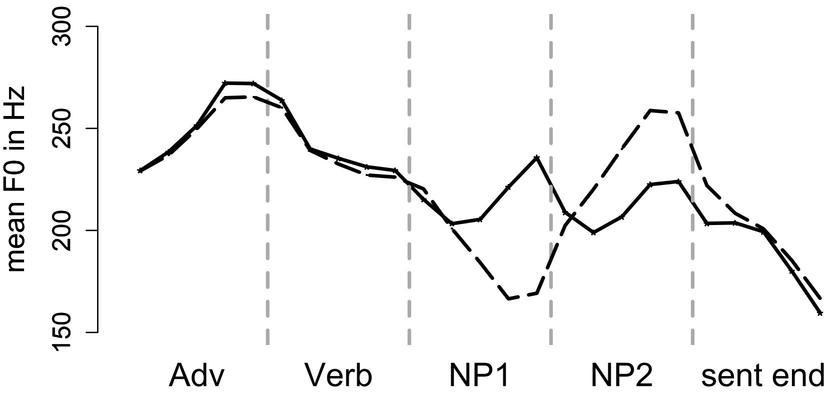

Figure 5.2: Mean pitch track, averaged over all speakers, of the adverbial phrase, the verb, the first NP and the second NP and the end of the sentence, for the genitive (solid line) and the dative condition (dashed line).

Constituent	Conditions	Means (f0max (Hz))	t-scores
ADV	GEN vs DAT	252 250	-0.779
VERB	GEN vs DAT	239 237	-1.653
NP1	GEN vs DAT	216 188	-14.45*
NP2	GEN vs DAT	212 235	13.11*

Table 5.2: Mean (f0max (Hz)) on four constituents of genitive and dative sentences as well as their statistical comparisons by means of t-tests.

The results of the prosodic analyses are summarized in Table 5.3 and make the prosodic difference on NP1 and NP2 among both conditions transparent: in the genitive condition, a rising pitch accent (L*H) is realized on NP1, signaling *continuation*. In contrast, a falling pitch accent followed by a low boundary tone (L%) is associated with the right edge of NP1 in the dative condition, signaling sentence *finality*. A second difference is evident on NP2 featuring a rising pitch accent (L*H) in both conditions, albeit with a significant difference in pitch accent scaling (cf. Figure 5.2).

Experiment	Condition	ADV-VERB	NP1	NP2
IP-boundary	GEN	L*H	L*H	L*H
	DAT	L*H	H*L L%	L*H

Table 5.3: Summary of the prosodic structure for both sentence conditions.

The analysis of duration

On the assumption that realizations before ι-boundaries tend to be longer than phrase medial realizations, there should be a difference in duration on the first NP between both conditions. The results of the durational analysis displayed longer durations on the first NP in the dative condition compared to the duration of the first NP in the genitive condition (cf. Table 5.4). Additionally, ι-boundaries are said to be followed by pauses. These pauses should result in a longer duration after the first NP in the dative condition compared to the pause duration after the first NP in the genitive condition (cf. Table 5.5). A linear mixed model with subject and item as random effects confirmed a significant effect of the durational difference of the first NP among both conditions. The dependent variable of this model is the constituent duration. Condition served as fixed effect, yielding a coefficient estimate of 66.36 with a standard error of 4.88 (t-value = 13.60) for the first NP. Equally, the analysis of the pause duration between NP1 and NP2 yielded a significant effect by fitting a linear mixed model with subject and item as random effects, pause duration as dependent variable and condition as fixed effect. The analysis resulted in a coefficient estimate of 64.40 with a standard error of 3.89 (t-value= 12.20). The phonetic analyses confirmed the prosodic difference between the dative and the genitive condition and thus validated a difference of the prosodic renderings of NP1 and NP2.

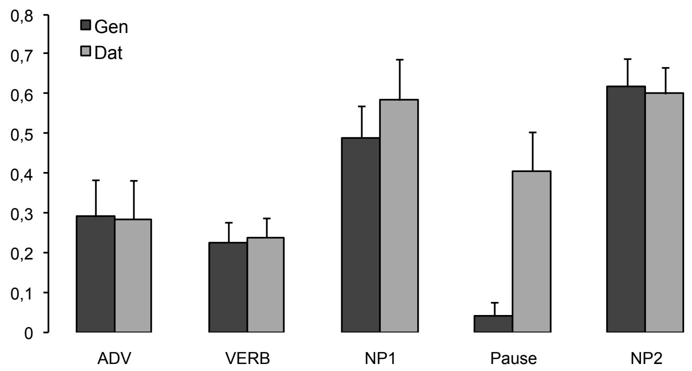

Figure 5.3: Mean duration times in s per constituent and pause across all speakers.

Constituent	Comparison	Means (duration (ms))	t-scores
ADV	GEN vs DAT	290 285	n.s.
VERB	GEN vs DAT	225 237	n.s.
NP1	GEN vs DAT	489 583	13.60
NP2	GEN vs DAT	619 600	n.s.

Table 5.4: Duration on the first four constituents of genitive and dative sentences as well as their statistical comparisons by means of t-tests.

Constituent	Comparison	Means (duration (ms))	t-scores
Pause 1	GEN vs DAT	39 406	12.20

Table 5.5: Pause duration between NP1 and NP2 of genitive and dative sentences as well as their statistical comparisons by means of t-tests.

Summary
The results suggest a difference in prosodic phrasing between both conditions which becomes evident on the first and second NP. The analyzed production data attest systematic variations of pitch accent type, pitch accent height and duration in the ambiguous region of the sentences. With respect to the genitive condition, the rising pitch accents on NP1 and NP2 that are approximately at the same height, as well as the results of the durational analyses revealing no *prefinal lengthening* at the position of NP1, suggest that NP1 and NP2 are phrased together within one φ-phrase in a recursive structure: (NP1 (NP2)φ)φ, where NP2 is embedded into NP1, as already suggested for the genitive condition presented in chapter 4.

In contrast, the dative condition shows a clear falling pitch accent followed by a low boundary tone on NP1 (H*L L%). Duration data revealed effects of *prefinal lengthening* on NP1, as well as a subsequent pause. Additionally, a rising pitch accent (L*H) on NP2 which returns to the tonal height comparable to the phrase-initial height of the adverbial phrase is evident. This tonal reset, as well as effects of *prefinal lengthening* and the low boundary tone can be taken as evidence for an ι-boundary at the end of NP1 (Truckenbrodt, 2005). Based on these phonetic properties, we propose the following prosodic phrase structure for the ambiguous region of the dative condition: (NP1)ι (NP2)φ.

5.2 Perception experiments

In the production experiment described above, we found differences in word duration, pause duration, pitch peak type and tonal phenomena such as downstep and upstep on ι-boundaries for the prosodic realization of the structures under scrutiny. Thus, the relevant difference in prosodic demarcation between the conditions presented in chapter 4 and the ones examined here in chapter 5 arises at the position of NP1 in the dative condition, featuring a bitonal falling pitch accent followed by a low boundary tone, while there is no such ι-boundary present in the dative condition presented in chapter 4. This tonal difference on NP1 between both sentence conditions examined in chapter 5 motivates the analysis of a further cue weighting, this time on ι-boundaries, to evaluate the perceptual relevance of f0 in comparison to word and pause duration. In order to make controlled assessment of the relative contribution of these cues for listeners, we carried out five perception experiments using stimuli with varying prosodic information to ascertain which of the prosodic cues listeners require (most) to disambiguate the structures. The logic is simple. If a certain prosodic cue is used for disambiguation, the absence (or inconclusiveness) of that information should have a detrimental effect on the consistency of listeners' ability to determine speakers' intended meaning in the forced-choice task. Such controlled perceptual studies allow us to (i) develop a prosodic cue weighting on ι-boundaries and (ii) to determine whether the prosodic correlates used for disambiguation are different from the ones used for smaller φ-phrase boundaries. These ι-boundary data, together with our findings of a prosodic cue weighting on φ-phrase boundaries eventually allow us to establish a correlation of the influence of each prosodic cue on listeners ability to assign one specific meaning of a syntactically ambiguous sentence, in dependence of the boundary size.

5.3 Experiment 1: Sentence disambiguation using original stimuli

Experiment 1 serves as baseline experiment evaluating if and to what extend listeners use the prosodic cues calculated in production in a sentence completion experiment. Based on the results of the sentence completion Experiment 1 presented in chapter 4 (cf. section 4.3), we hypothesize for the present Experiment 1 that prosodic correlates of the speech signal will influence listeners' computation of the underlying syntactic structure. An effect of the prosodic structure is anticipated as follows: the prosodic correlates associated with the genitive condition guide listeners' disambiguation in such a way that they disambiguate the auditory sequence more often towards the genitive interpretation in a forced choice sentence completion experiment. In case of the dative condition, the corresponding prosodic correlates guide participants more often to opt for the dative sentence continuation as opposed to the genitive one.

Material

16 sentences of each condition were pronounced by a phonetically trained speaker according to the intonational patterns measured in the production study. A plotted pitch contour of each sentence, superimposed for all sentences per condition confirmed the congruence in prosodic structure of these recordings with the prosodic contour of the production data for all words in the temporarily ambiguous region (cf. Figure 5.4 below). In detail, a rising pitch accent (L*H) was realized on the adverbial-verb phrase in both conditions. For the following constituent NP1, a rising pitch accent (L*H) was pronounced for the genitive condition, while a falling (H*L L%) accent was realized for the dative condition, respectively. The remaining contour was tonally identical with a rising pitch accent (L*H) on NP2 in both conditions, albeit with different scaling properties of the rising accent, which was at a comparable height relative to the preceding rising accent on NP1 in the genitive condition, and upstepped comparable to the height of the phrase initial accent in the dative condition.

The bitonal rising pitch accent on NP1 in the genitive condition, and the bitonal falling pitch accent on NP1 in the dative condition constitute two different accents, appearing as a salient perceptual feature easily attributable to either of the two sentence conditions. Therefore, two further experimental conditions have been devised with pitch accents on NP1 identical to the ones reported for the genitive and dative condition, respectively. Specifically, a second dative condition with a rising accent on NP1 followed by a high boundary tone (L*H H%) was created, as well as a second genitive condition with a falling accent on NP1 (H*L)[1], to complement the data set. The additional conditions were recorded by the phonetically trained speaker who spoke the stimuli of the original conditions. The two additional conditions will be referred to as «non-original conditions» in the remaining part of this work. Table 5.6 summarizes the entire data set and Figure 5.4 depicts the plotted f0-contours for all four experimental conditions. The non-original experimental conditions had been devised to control the accent pattern on NP1 in such a way that one particular accent is not necessarily associated with one of the two sentence structures in order to exclude response strategies. Each experimental sentence of the non-original conditions was uttered according to the phonetic requirements. Sentences that did not match the mean intonational contour of the production experiment were replaced by new recordings. In this way, a consistent data set was generated by one speaker. Subsequently, the disambiguating part of all target sentences (i.e. the contextual information) was cut off. Each fragment was equated in loudness at the level of 70 db using a scale intensity praat script. The resulting sentence fragments were taken to test listeners' use of prosodic information to predict upcoming entities in a forced-choice task to eventually examine a prosodic cue weighting on ι-boundaries.

[1] Prenuclear accents can be realized with a falling H*L contour (Féry, 1993).

Experiment	Condition	ADV-VERB	NP1	NP2
ι-boundary	GEN-nat	L*H	L*H	L*H
	DAT-con	L*H	L*H H%	L*H
	DAT-nat	L*H	H*L L%	L*H
	GEN-fall	L*H	H*L	L*H

Table 5.6: Summary of the experimental conditions and their prosodic structure

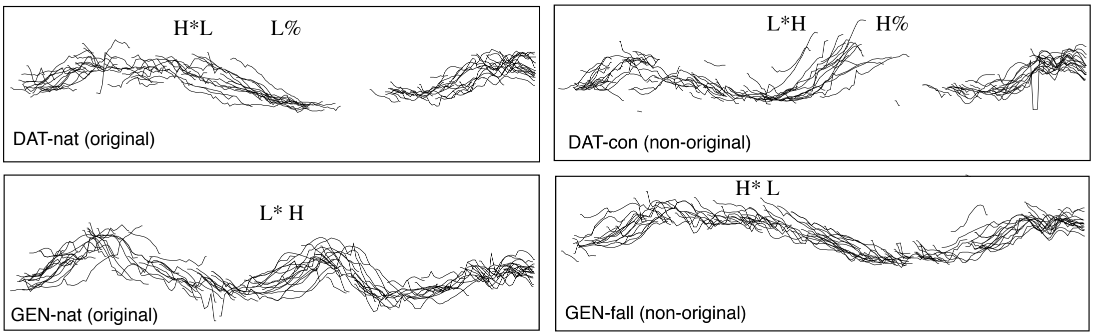

Figure 5.4: Plotted f0-contours for all experimental conditions, averaged over all sentences per condition. The upper panel shows the f0 contour of the dative conditions (original f0 contour on NP1: H*L L% (left) and non-original f0 contour on NP1: L* H H% (right). The lower panel shows the genitive condition (original f0 contour on NP1: L*H (left) and non-original f0 contour on NP1: H*L (right).

Method

Procedure

The experimental sentences (16 x 4 = 64) were allocated to four lists using a Latin square design[2]. The latin square design was used such that each participant could only see 4 different sentences from each of the 4 sets. In this way, each participant was presented with 16 experimental sentences, 4 from each condition. The 16 experimental sentences were embedded in 16 filler sentences from two unrelated experiments to discourage response strategies. As a result, each experimental list contained 32 items. In a forced-choice task identical to the one described in the previous chapter 4, participants listened to each sentence fragment up to and including NP2, and then were asked to complete the sequence by choosing the better fitting of the two offered context continuations yielding either to the genitive or the dative interpretation of the second NP. Listeners were aware that both choices were plausible but were forced to choose the most appropriate response depending on how the sequence was uttered. Participants performed a 3-item training before the experiment began. All of the judgements of this and of all following experiments were off-line, meaning that subjects were not under time pressure to make their choices.

Participants

56 undergraduate students from the University of Potsdam participated in the experiment. They were native speakers of German and reported no speech or hearing impairments. The participants did not take

[2] The increased number of experimental sentences (n = 64) compared to the data set presented in chapter 4 allowed a Latin square design.

part in any of the previously presented experiments and were naive as to the purpose of the experiment. They either received course credits or were paid for participation.

Results

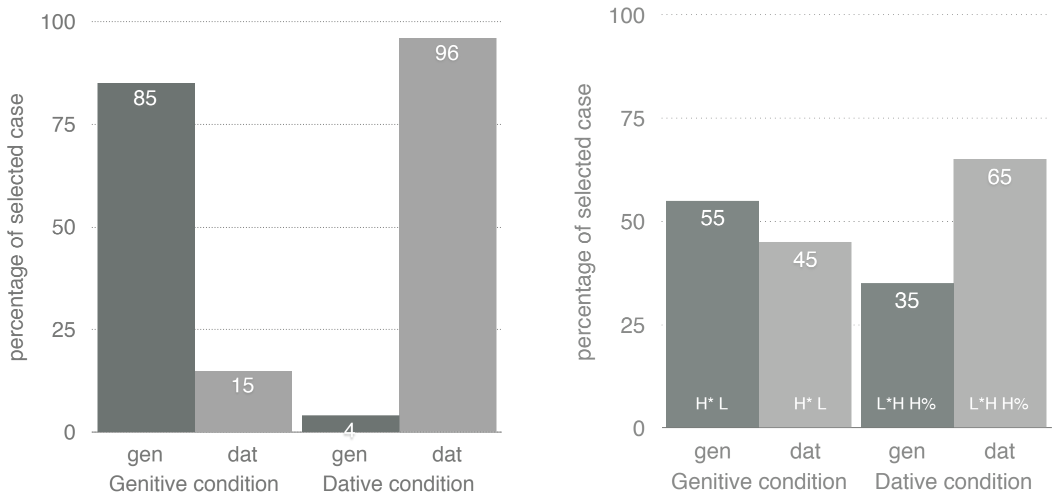

Figure 5.5: Percentage of the selected sentence continuation in Experiment 1, for the original sentences (left) and for the non-original conditions with inverted boundary tones (right).

Figure 5.5 depicts the response pattern of Experiment 1 for the two original conditions (left graphic) and the two non-original conditions bearing a reversed pitch accent on NP1 (right graphic). The results are very clear in that listeners chose the appropriate sentence continuation on prosodic grounds. For the original conditions, in 85% of the responses, the correct genitive continuation was selected for a presented genitive sequence, and for a presented dative sequence, in 96% of the cases, the prosodically fitting dative continuation was picked by listeners. For the non-original conditions, subjects' performance deteriorated significantly: for the non-original genitive condition, listeners selected in 55% of the cases, the genitive version out of the two alternative sentence continuations, while in 65% of the cases, the dative sentence continuation for the non-original dative condition was selected. A statistical frequency-based analysis, using crossed random factors subject and item, was computed for all conditions. Sentence sequence (GEN, DAT) served as fixed factor. The statistical analysis relied on the sentence continuation as dependent variable yielding a coefficient estimate of -2.361, with a standard error of 0.534 (z-value = -4.42), confirming the occurrence of a clear disambiguating effect of the original conditions. In contrast, the identical statistical analysis for the non-original sentence conditions yielded a coefficient estimate of 0.508, with a standard error of 0.20 (z-value = 1.94), revealing a missing disambiguating effect since the selection of the sentence continuation was non-significant. Juxtaposing the results of the original and non-original conditions indicate that the quality of the pitch accent (rising or falling) at the position of NP1 affects listeners' disambiguation. As proposed in Féry et al. (2009) and confirmed by the results of Experiment 1, tones are attributed meaning and therefore influence listeners' anticipations of as yet unperceived sentence material. Transferred to the sentence material of the non-original conditions, the perception results suggest that the rising accent of the non-original dative condition is interpreted as non-final, while the falling accent of the non-original genitive condition is interpreted as final. This expectation clashes with the actual sentence continuation in both conditions and induces the phonological violation responsible for the decrease in disambiguation of the non-original conditions. To summarize, the results establish that the pitch accent type at the position of NP1 is necessary (and possibly sufficient) to compute the correct sentence interpretation, and thus selecting the correct sentence continuation.

For the following four perception experiments, we have endeavored to approximate the prosodic conditions under which disambiguation is actually done by having listeners do the same sentence completion task described in the previous Experiment 1 but this time with manipulated speech. This way we can tease apart the various prosodic cues which have been identified in the production experiment to be relevant for disambiguation in order to establish a prosodic cue weighting associated with ι-boundaries. Below, we report four perception experiments, one for each performed manipulation. The reason we ran one experiment for each manipulation instead of one experiment containing all conditions was that it would have complicated the design considerably, leading to a much smaller number of trials in every stimulus/condition design cell. Instead we wanted to collect enough data to derive individual response distributions for every stimulus in every condition and for every performed manipulation, in parallel to the perception experiments presented in chapter 4.

5.4 Experiment 2: Influence of the intonational contour

The aim of Experiment 2 was to determine whether listeners would make use of the intonational contour to successfully disambiguate both sentence structures. To accomplish this, the sentence fragments employed for Experiment 1 were prosodically manipulated by flattening the contour at an equal level in all four experimental conditions, this way the remaining prosodic information of the duration (i.e. *prefinal lengthening* on NP1 and a pause between NP1 and NP2) was the only decisive cue listeners could rely on when processing the sentence structures. Since the production experiment evidenced a significant difference of the prosodic contour on NP1 among both sentence condition, we hypothesize that listeners' fail to disambiguate the two sentence conditions with an intonational contour entirely removed.

Material

Flattening the intonational contour was done using the PSOLA resynthesis function in praat (Boersma & Weenink, 2013). The pitch cues of the original fragments were set to a constant value of 120 Hz, such that the pitch contour was completely horizontal (cf. Figure 5.6, right graphic).

Method

Procedure
The experimental procedure was identical to those used in Experiment 1, except that listeners were informed that the sentence material was manipulated acoustically. The same latin square design described

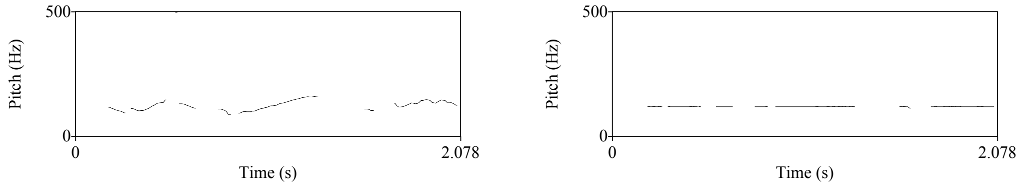

Figure 5.6: Illustration of an unmanipulated stimuli (left) and a f0-manipulated stimuli (right). The graphics picture the ambiguous sequence up to and including NP2

in 5.3.2 was used in the current Experiment 2. After three practice trials, the experiment started.

Participants
56 undergraduate students from the University of Potsdam participated in the experiment. They were native speakers of German and reported no speech or hearing impairment. The students were naive as to the purpose of the experiment and received course credit for participation. Participation in any of the previously reported experiments was excluded.

Results

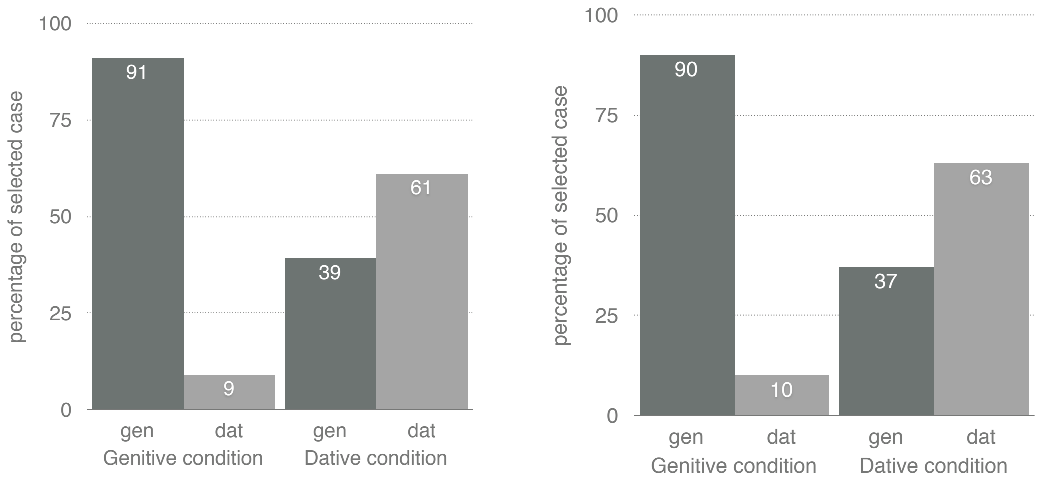

Figure 5.7: Percentage of the selected sentence continuation in Experiment 2, for the original sentences (left) and for the non-original conditions with inverted boundary tones (right).

Figure 5.7 shows the response pattern of Experiment 2 for the two original conditions (left graphic) and for the two non-original conditions bearing a reversed pitch accent on NP1 (right graphic). For the sake of completeness, both sentence conditions are reported below and pictured in Figure 5.7, though the performed prosodic manipulation of a flattened intonational contour eliminated the prosodic difference between the original conditions and the non-original conditions. Testing both sentence conditions in the current experiment as opposed to the original conditions only, was to allow us to use the same latin square design introduced in Experiment 1 above.

With an intonational contour entirely removed (leaving intact word and pause duration), subjects are still able to successfully perform the sentence completion task for the original conditions (left graphic). In detail, in 91% of the cases, the correct genitive continuation was selected when listeners heard the ambiguous truncated sentences carrying durational properties of the genitive condition, while in 61% of the cases, the dative sentence continuation was correctly chosen for the ambiguous truncated sentences carrying durational properties of the dative condition.

Very similar to the original conditions, and certainly not surprisingly, the removal of the intonational contour in the non-original conditions did not change subjects' performance in the sentence completion task (right graphic). In comparison, in 90% of the responses, the genitive verb was selected for a presented genitive sequence, and in 60% of cases, listeners opted for the dative sentence continuation when they listened to a dative sequence. We performed a statistical frequency-based analysis using crossed random factors SUBJECT and ITEM for all four conditions. Sentence sequence (GEN, DAT) served as fixed factor. The statistical analysis relied on the sentence continuation as dependent variable yielding a coefficient estimate of 3.653, with a standard error of 0.406 (z-value = 8.99), confirming the occurrence of a clear disambiguating effect of the original conditions. In contrast, the identical statistical analysis for the non-original sentence conditions yielded a coefficient estimate of 4.139, with a standard error of 0.605 (z-value = 6.83), revealing the same clear disambiguating effect. The finding that the removal of the f0 contour still allows listeners to disambiguate both sentence structures does not confirm our hypothesis for Experiment 2, and rather suggest that when no decisive intonational information is present (no information of the pitch accent is available), the remaining information of duration is called into play (either the difference in pause duration between NP1 and NP2, or the difference in *prefinal lengthening* on NP1).

The influence of the *pause duration* will be attested in Experiment 3. Separately, the influence of the prosodic cue *prefinal lengthening* will be examined in Experiment 4.

5.5 Experiment 3: Influence of pause duration

The production study revealed pause duration between the first and second NP as a highly significant prosodic difference between the two ambiguous sentence structures. Thus, the objective of Experiment 3 was to investigate whether listeners use the information of the pause duration, i.e. the duration between the first and second NP as a decisive prosodic cue to disambiguate the syntactic structures prosodically. A manipulation of the pause duration is therefore necessary to gauge the influence of the particular prosodic information on disambiguation. We hypothesize for Experiment 2 that with manipulated pause duration the ability to ascertain the meaning of the two condition decreases.

Material

The sentence material used for Experiment 3 was prosodically manipulated so that the pause duration between NP1 and NP2 of each ambiguous sentence pair carried the pause duration of its alternate condition by exchanging the pause duration of one condition with the value of the pause duration of the alternative condition. The pause duration for each ambiguous sentence pair was measured and the corresponding values from one condition were inserted into the other condition, respectively. Modifications of the pause durations to the stimuli for each sentence pair were done using praat (Boersma & Weenink, 2013). Table 5.7 illustrates the results of the manipulation procedure and shows the original pause duration and the inserted pause duration value for all four sentence conditions, illustrated for two sentence pairs (P stands for the pause between the first and second NP). This pause duration manipulation resulted for the dative conditions to carry the pause duration of the alternative genitive conditions, exemplified for 2 of the 16 experimental sentences (cf. Table 5.7). Appendix C contains a table showing the pause manipulation for the entire stimulus set used in Experiment 3.

manipulated pause	condition	original duration	compressed/ expanded to
1. |NP1 | P |NP2|	DAT-nat	388	000
	GEN-nat	000	388
	DAT-con	147	054
	GEN-fall	054	147
2. |NP1| P |NP2|	DAT-nat	340	001
	GEN-nat	001	340
	DAT-con	272	039
	GEN-fall	039	272

Table 5.7: Illustration of the pause duration manipulation in ms for the original conditions (DAT-nat and GEN-nat) and for the non-original conditions (DAT-con and GEN-fall) for two sentence pairs.

Method

Procedure and Participants

The procedure and experimental settings were identical to previous experiments. The experimental sentences were arranged in the same latin square design described above for Experiment 1 and 2.

56 undergraduate students from the University of Potsdam participated in the experiment. All participants were native speakers of German and reported no speech or hearing impairments. They did not participate in any of the previously reported experiments and were naive as to the purpose of the experiment. Their participation was compensated by course credit assignments.

Results

Figure 5.8 presents the results of the sentence completion experiment. For the original sentence structures that carried a natural intonational contour but manipulated pause durations between the first and second NP, disambiguation was still achieved but on a lower level compared to the unmanipulated version in Experiment 1.

In detail the results show, that when an original genitive carries a dative pause duration between NP1 and NP2 (which corresponds to an ι-boundary pause), the selection of the genitive sentence continuation drops from 85% (Experiment 1, left graphic) to 62% (Experiment 3, left graphic). Similarly, for a dative structure carrying a genitive pause duration between NP1 and NP2, the choice for selecting a dative sentence continuation significantly drops from 96% (Experiment 1, left graphic) to 70% (Experiment 3, left graphic). However, this pause manipulation still allows listeners to compute the correct interpretation for the original conditions (Figure 5.8, left graphic). The corresponding statistical frequency-based analysis, using crossed random factors subject and item, and sentence condition (GEN, DAT) as fixed factor was computed for both original conditions; sentence continuation served as dependent variable yielding a coefficient estimate of -1.476, with a standard error of 0.3307 (z-value = -4.465), confirming the occurrence of a clear disambiguating effect for the pause manipulated original conditions.

In those cases where the same pause manipulation was performed on the non-original conditions, disambiguation was not possible anymore (Figure 5.8, right graphic). That is, it was possible to completely shift the meaning of an expression uttered with one meaning into its alternate meaning when two prosodic cues (pause duration and the pitch accent type at the position on NP1) work against the original condition. Thus, the effect of pause duration and pitch contour were additive in total proportion and led listeners towards the wrong sentence interpretation. The statistical analysis for the non-original sentence conditions yielded a coefficient estimate of 1.3429, with a standard error of 0.3058 (z-value = 4.392), revealing the same significant difference in response as observed for the original conditions, albeit towards the opposite direction.

To summarize, the results of Experiment 3 support the findings of Experiment 1 and 2 in showing that the boundary tone preceding the phrase boundary seems to be a crucial source of prosodic information in order to correctly parse the sentences, and that parsing is still possible even in the case of manipulated pause duration within the ambiguous region.

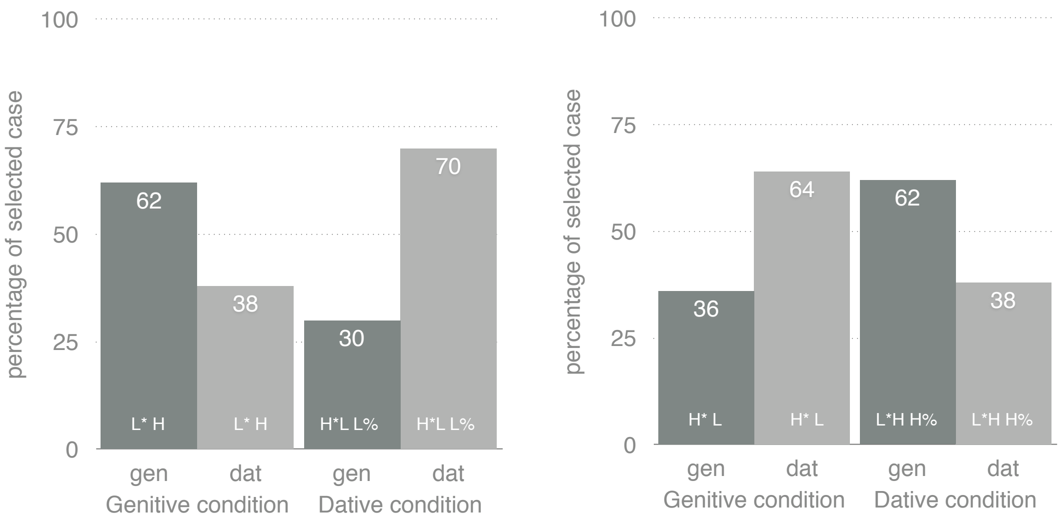

Figure 5.8: Percentage of the selected sentence continuation in Experiment 3. The left graphic displays the results for the original sentences; the right graphic shows the results for the non-original conditions with inverted boundary tones.

5.6 Experiment 4: Influence of prefinal lengthening

To examine the effect of *prefinal lengthening* on listeners' disambiguation, a perception study was carried out that used sentence material with manipulated durational information of NP1. This manipulation was done independently from pause duration, i.e. the stimuli of the original condition contained natural pause and f0 information, but manipulated word duration to disentangle the exclusive impact of *prefinal lengthening* as opposed to *pause duration* on ambiguity resolution.

The scope of prefinal lengthening and its determining variables have been examined in a number of studies, see Turk & Shattuck-Hufnagel (2007) for a review. Generally, it has been recognized that prefinal lengthening correlates with the size of the following phrase boundary (Price et al., 1991; Shattuck-Hufnagel & Turk, 1996) and the degree of prefinal lengthening reduces with the distance towards the phrase boundary (Byrd et al., 2006). More specifically, prefinal lengthening affects all unstressed syllables up to and including the first most prominent syllable left adjacent to the phrase boundary (Kohler, 1983; Berkovits, 1994; Turk & White, 1999). Turk & Shattuck-Hufnagel (2007) additionally found that lengthening is progressive - syllables affected from prefinal lengthening show stronger lengthening effects in the coda than in the nucleus, and stronger effects in the nucleus that in the onset. According to this, the durational properties of both syllables of NP1 have been manipulated in the current experiment 4.

Since the prosodic difference of *prefinal lengthening* on NP1 among both conditions was not as large as the difference in *pause duration* between NP1 and NP2, we expect less of an influence of *prefinal lengthening* on disambiguation, and hypothesize that the original conditions are disambiguated on a similar level as it was the case for the prosodically unmanipulated original structures shown in Experiment 1. Likewise we hypothesize for the non-original conditions that listeners fail to disambiguate the structures on similar grounds.

Material

The sentence material used in Experiment 1 was manipulated as follows. For each ambiguous sentence pair, the duration of NP1 was measured. Then, the duration of NP1 had been altered by adjusting its duration in the genitive condition with the duration of NP1 in the dative condition, and vise versa. As a result, each manipulated NP1 carried the duration properties of its alternative condition. The duration of NP2 remained untouched for two reasons: the duration differences of NP2 between both conditions were not significant and second, the current experiment aimed to identify an influence of *prefinal lengthening* only. Table 5.8 shows the original duration and the manipulated values for all four conditions (original and non-original) exemplified for 2 of the 16 experimental sentence pairs. The first example is referred to as 1. |NP1|, and the second example is referred to as 2. |NP1|, cf. Table 5.8. The manipulation for the entire stimulus set used in Experiment 4 is given in Appendix C.

Method

Procedure and Participants
The procedure and experimental design were identical to previous experiments. 56 undergraduate students from the University of Potsdam participated in the experiment. They were native speakers of Ger-

manipulated words	condition	original duration	compressed/ expanded to
1. |NP1|	DAT-nat	594	511
	GEN-nat	511	594
	DAT-con	639	616
	GEN-fall	616	639
2. |NP1|	DAT-nat	624	465
	GEN-nat	465	624
	DAT-con	619	565
	GEN-fall	565	619

Table 5.8: Word duration manipulation in ms of genitive and dative sentences shown for two sentence pairs.

man and reported no speech or hearing impairment. Subjects did not participate in previously reported experiments, were naive as to the purpose of the experiment and received course credit for participation.

Results

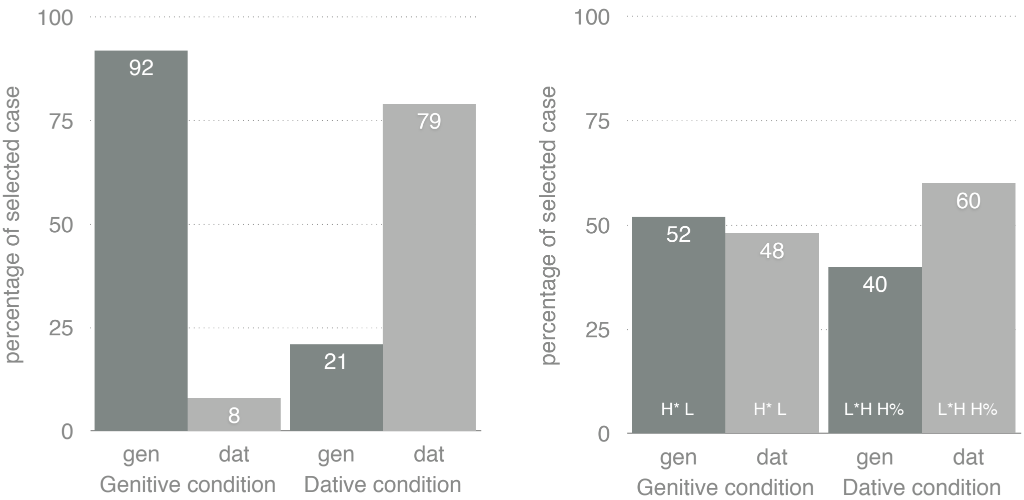

Figure 5.9: Percentage of the selected sentence continuation in Experiment 4. The left graphic displays the results for the original sentences; the right graphic shows the results for the non-originial conditions with inverted boundary tones.

For the original stimuli pictured on the left graphic in Figure 5.9 we see that participants were highly skilled at disambiguating the two versions of the syntactically ambiguous sentences. For the dative condition, listeners selected the dative sentence continuation in 79% of the cases and, for the genitive condition, participants accurately selected the genitive sentence continuation in 92% of the cases. Given that the disambiguation ability for the unmanipulated original sentences was similarly high in Experiment 1 (Experiment 1: dative condition: 96% dative sentence continuation, genitive condition: 85%

genitive sentence continuation), this is remarkably similar and reflects what is found for the sentence material carrying natural prosodic correlates.

In Figure 5.9 on the right, the response pattern for the manipulated non-original sentences, i.e. the sentence conditions carrying a reversed boundary tone at the position of the first NP, is displayed. The reversed boundary tone information in combination with the exchanged information of the *prefinal lengthening* on NP1 provided a prosodic combination which disallowed listeners to correctly deduct the hidden structure. Instead listeners performed on chance level during sentence parsing. The parsing results were confirmed by statistical analyses. Fixed factors, random factors and dependent variables were identical to the ones reported for the previously presented statistics revealing for the original conditions a coefficient estimate of -1.631, with a standard error of 0.2699 (z-value = -3.981), confirming the occurrence of a clear disambiguating effect. The identical statistical analysis for the non-original sentence conditions yielded a coefficient estimate of 0.308, with a standard error of 0.40 (z-value = 0.92), revealing a missing disambiguating effect since the selection of the sentence continuation was non-significant.

The most important finding derived from Experiment 4 is that the exchange of *prefinal lengthening* on the first NP did not affect the disambiguation substantially. In fact, the performed durational manipulation on the original and non-original conditions led to the same response pattern as was the case for the unmanipulated sentence conditions presented in Experiment 1 above. Therefore, the result of Experiment 4 can be taken as a further piece of evidence that the information of the boundary tone at the position of NP1 seems to be a relevant prosodic correlate. In contrast, the information of the *prefinal lengthening* had no effect on listeners' disambiguation ability.

5.7 Experiment 5: Influence of downstep and upstep

Alongside durational cues and boundary tone realizations, upstep and downstep phenomena have been identified in the production study reported in 5.1 as being a diverse prosodic rendering of the genitive and dative condition. While an upstep on the second NP in the dative condition can be taken as evidence for the formation of a new prosodic phrase, the same NP in the genitive condition appears to be downstepped towards the preceding one, suggesting the first and second NP to be phrased together. In a perception study, Féry and colleagues found an influence of such tonal cues on listeners' expectation on not-yet-perceived constituents. Féry et al. (2009) investigated tonal contours of truncated sentences in German and showed that tonal properties seem to trigger the anticipation (in that case the information structure) of upcoming sentence material. The current and last experiment of this dissertation investigates whether the observed tonal difference, here referred to as upstep and downstep, similarly trigger anticipatory effects and may bias the resolution of the syntactic ambiguity.

Material

The acoustic analysis of the speech material revealed a systematic prosodic difference on the second NP (i.e. upstepped rising accent vs. downstepped rising accent) between the two conditions, suggesting a specific phrasing dependent on the syntactic structure. To scrutinize the influence of the pitch accent height of NP2 on listeners' ability to resolve the syntactic ambiguity, two novel sentence conditions were derived by cross-splicing the first part (up to and including NP1) of the genitive condition together with the second part (NP2) of the dative condition, and vise versa. The splicing procedure in combination with an amplitude normalization at 70 db protected against detectability of the speech manipulation at the splicing point. Figure 5.10 depicts the plotted f0-contours for the cross-spliced conditions (right) and the original f0 plot for comparison (left). Appendix C lists the exact f0 maxima on NP2.

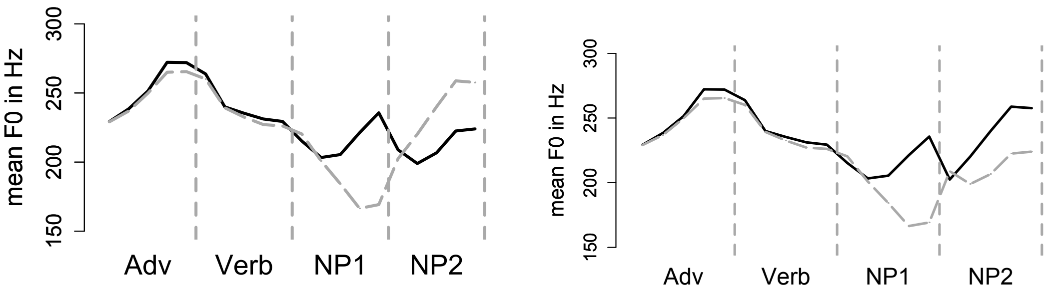

Figure 5.10: Stylized pitch contours from the beginning of the sentence up to and including the ambiguous NP2, broken down by experimental condition. The solid black line presents the genitive condition, the dashed grey line pictures the dative condition. The cross-spliced conditions with the splicing point after NP1 are shown on the right and, as a comparison, the original f0 plot is pictured on the left.

If the relative height of the pitch accent on NP2 is a crucial prosodic cue for triggering the correct sentence continuation of truncated sentences and thus resolves the syntactic ambiguity, we hypothesize that in the upstepped rise variant (genitive condition with a high rise), the upstepped accent on NP2 would be interpreted as the rising accent on the subject of the subsequent sentence, which in turn, must be followed by a dative sentence continuation. We therefore expect participants to complete the sentence fragments more often with the dative sentence continuation in this condition. On the other hand, in the variant with a downstepped rise (dative condition with a low rise), the contour on the second NP should be interpreted as a rising accent of the genitive modifier. In this case, we expect participants to complete the sentence fragments more often with the genitive continuation.

Note, that the non-original stimulus conditions were not tested in Experiment 5, as the focus of inspection was the upstepped vs. downstepped rise on NP2 only. The non-original conditions with an additional prosodic manipulation of the rise on NP2 would have confounded the results.

Method

Procedure and Participants

The experimental procedure and design were identical to those used in all previous experiments. 56 undergraduate students from the University of Potsdam were recruited to participate. All were native speakers of German and reported no speech or hearing impairment. The participants were naive as to the purpose of the experiment and received either credits towards completing a course requirement or financial reimbursement for participation. None of the subjects participated in either of the previously presented production and perception experiments.

Results

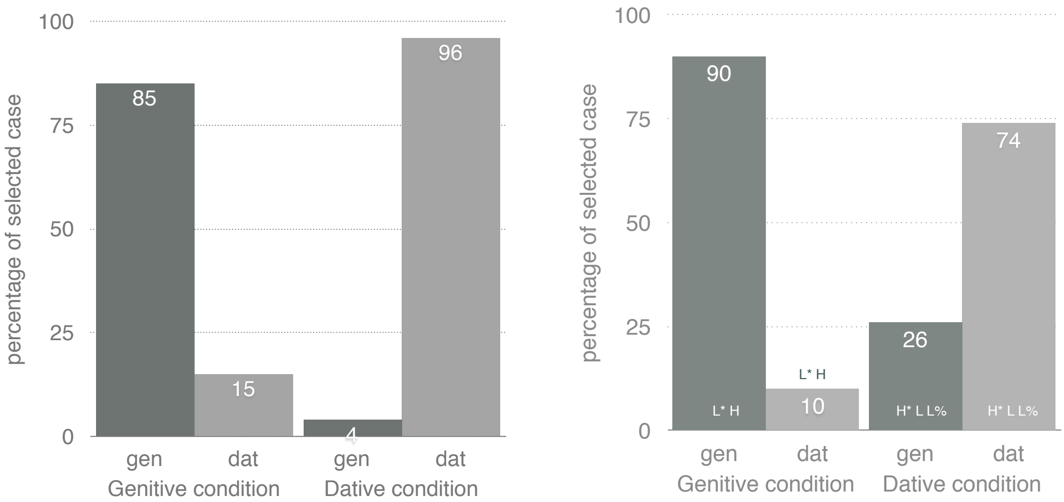

Figure 5.11: Percentage of the selected sentence continuation in Experiment 5 (right graphic), displaying the results for the cross-spliced sentences, compared to Experiment 1 (left graphic), presenting the results of the forced-choice experiment using unmanipulated stimuli.

The results of the sentence completion experiment using cross-spliced stimuli are displayed in Figure 5.11 (right graphic) and show a clear tendency in favor of the given sentence condition: in the upstepped rise variant (i.e. the genitive condition concatenated with a high rising accent on NP2), listeners selected in 90% of the examples the genitive sentence continuation. Similarly for the downstepped rise condition (i.e. the dative condition concatenated with a low rising accent on NP2), the dative sentence continuation was selected in 74% of the cases. The statistical frequency-based analysis using crossed random factors subject and item was computed for both conditions. For the statistical analysis, sentence sequence (GEN, DAT) served as fixed factor, the verb served as dependent variable yielding a coefficient estimate of 3.88, with a standard error of 0.371 (z-value = 10.43), confirming the occurrence of a clear disambiguating effect of the cross-spliced conditions. Comparing the results of Experiment 5 with the outcome of Experiment 1 where no prosodic manipulation of the sentence material was performed, a similar result was obtained (cf. Figure 5.11 left graphic, and section 5.3 for the original conditions and its statistical analysis). Therefore, we conclude, that the preference for selecting the correct sentence continuation is not affected / influenced by the height of the rising pitch accent on NP2: participants were equally able to choose more dative than genitive sentence continuation when they heard a dative sequence with a low rise on NP2, and similarly, they selected more genitive than dative sentence continuations when faced with a genitive sequence featuring an upstepped rise on NP2. This supports our findings of Experiment 1 and 2 by revealing that the boundary tone continues to be relevant and obviously the most important prosodic cue upon which listeners base their syntactic decisions. These results do not confirm the possibility that the height of the pitch accent on the second NP is a prosodic determinant for the response distributions, and therefore we conclude that listeners do not draw on that particular prosodic information when disambiguating the syntactic ambiguity under scruitiny. Rather, the results suggest that (i) the point of disambiguation takes place before processing NP2, and (ii) that the prosodic demarcation of the first NP (i.e. rising vs. falling accent) is the crucial disambiguating prosodic information.

5.8 Summary of the experimental findings in support of the boundary tone as disambiguating cue

Our experimental set-up for the five experiments involved the manipulation of the speech signal to test the individual influence of each prosodic cue on sentence disambiguation to establish a cue weighting on ι-boundaries. It appeared that the experimental paradigm was tapping into the cognitive process of disambiguating syntactically ambiguous structures. The finding of our perception experiments using different versions of prosodically manipulated truncated sentences demonstrated that indeed this was the case. The corresponding perception experiments on ι-boundaries are summarized in the following section.

First, for the natural (original) stimuli in **Experiment 1**, result patterns were very distinct. Subjects were highly skilled at disambiguating the truncated sentences by prosodic means. On the other hand, the results of the sentence completion study for the non-original conditions showed that the prosodic manipulation (i.e. inverting the boundary tone preceding the prosodic boundary from low to high and vise versa) influenced processing in such a way that listeners were no longer able to distinguish the two sentence structures. The conclusion can be drawn that the prosodic information of the intonational contour is taken into account while parsing the sentence structures. More specifically, the findings of a clear disambiguating effect in the original conditions and a missing disambiguating effect in the non-original conditions speaks in favor of the boundary tone (low vs. high) at the position of the first NP to be the crucial source of information.

Intriguingly, the results of **Experiment 2** showed that the removal of the f0 contour (leaving words and syntax intact) had only partial influence on listeners' disambiguation ability. Though it led to a slightly degraded discrimination ability for the dative condition (in comparison to Experiment 1), disambiguation was still possible without intonational cues when original pause duration and prefinal lengthening information were kept. This result suggests two things. Firstly, the findings support the assumption that it is not the intonational contour in general, but the correct phonological contrast of a low vs. high boundary tone that listeners need in order to successfully interpret the sentences. Secondly, under the absence of f0 information, the remaining durational properties offer sufficient prosodic information to correctly distinguish the ambiguous sentences.

Experiment 3 allowed us to make predictions as to the observed influence of durational correlates. By replacing the pause duration between the first and second NP in one sentence condition with the corresponding value of the other condition, disambiguation was still achieved but on a lower level compared to Experiment 1. Two further conditions emerged from manipulating the boundary tone on NP1 into the opposite direction (we referred to these conditions as non-original conditions). For these non-original conditions we showed that when the same pause manipulation was performed, successful disambiguation was no longer possible. Rather, the effect of pause duration and pitch contour were additive in total proportion and led listeners towards the wrong sentence interpretation.

Experiment 4 revealed that when the prefinal lengthening information on NP1 were entirely replaced with the prefinal lengthening values of the opposite structure and vise versa, there was no change in subjects' disambiguation ability - they were almost as accurate as with the original (unmanipulated) recordings in Experiment 1. In what follows, the effect of prefinal lengthening does not, on its own, contribute to disambiguation when the remaining prosodic cue of the pause duration and boundary tone information are present. If the same prefinal lengthening manipulation is performed for the non-original conditions, such that two prosodic cues *prefinal lengthening* and the boundary tone information are reversed, listeners' performance deteriorated significantly and disambiguation was no longer possible. These results give further evidence that the quality of the boundary tone at the position of the first NP is necessary for determining the correct syntactic structure when processing ι-phrase boundaries.

Experiment 5 tested an influence of the pitch accent scaling of NP2 on disambiguation by presenting cross-spliced versions of the original sentences. The experiment revealed that when hearing the truncated cross-spliced versions, subjects could indeed anticipate the correct sentence continuation, with the same high degree of accuracy found for the unmanipulated sentences (Experiment 1). This was interpreted as evidence that listeners weight tonal cues on ι-boundaries more heavily than tonal properties following ι-boundaries. To summarize, when combining the results of Experiment 5 with the results obtained in Experiment 1, it becomes evident that the different scaling of the L*H accent on NP2 (i.e. an upstep in the dative condition with an averaged mean value of 235 Hz and a downstepped accent in the genitive condition with an averaged mean value of 212 Hz) does not change the result pattern. Rather, the results of Experiment 5 support the findings from Experiment 1 in that listeners are well capable of correctly disambiguating the structures as long as meaningful boundary tone information on NP1 are available. A certain degree of intonational flexibility on NP2 does not influence the result pattern obtained for the unmanipulated sentence version.

5.9 General discussion

Psycholinguistic research is mainly dominated by production studies as opposed to perception studies. This situation results from the comparatively poor access to the processes underlying perception compared to the opportunities offered by experimental work on speech production. Further, research in perception requires an increased approach on the experimental method. With our experimental design in the present thesis we offered a method to directly compare production with perception data and examined how prosodic information in the auditory signal affect sentence comprehension in a forced-choice paradigm. The experimental method of a forced-choice paradigm, a design in which two possible choices were presented as answers to each trial appeared to be highly valuable. It minimizes biases as subjects were aware that either one of the two presented choices was equally possible, compared to discrimination tasks for example. Additionally, it is suitable for very similar stimuli, and results in frequency distributions easily accessible. Forced-choice experiments are generally considered to be comfortable for participants as well (Macmillan & Creelman, 2005). The method proved to be highly sensitive towards the prosodic manipulation and allowed us to establish a prosodic cue weighting on φ- and ι-boundaries.

In particular, the experiments presented in this thesis examined the processing of case ambiguous structures which also provided a powerful, focused test that seemed to be suitable for eliciting a prosodic cue weighting associated with φ- and ι-boundaries in German. From the nature of the constructions used to this end, essentially two interesting findings were obtained. First, the results show that not only are German listeners sensitive to phrasal and clausal prosody, but they make use of it to discriminate prosodic phrase boundaries. Further, results reveal that German listeners are sensitive to all three correlates of prosodic phrase boundaries (f0, pause duration and prefinal lengthening) that the manipulation of each of these cues affects listeners' discrimination ability of two ambiguous sentences in a sentence completion experiment. Thereby, listeners discrimination ability is influenced by two variables, (i) by the size of the prosodic boundary, and (ii) by the prosodic informational source of the speech signal. In the following sections, the main results will be discussed in turn, before a model of phonological perception in an Optimality-Theoretic framework is considered in the next chapter.

Flattened intonational contour

The removal of the pitch contour from the original sentence material did not have an effect on subjects' sentence completion performance. This could mean that in our experimental setting with ambiguous sequences containing ι-phrase boundaries, listeners made recourse to the remaining durational cues in order to ascertain the meaning of the ambiguous sentence material[3]. That the presence of only a limited set of prosodic boundary cues still allows listeners to successfully disambiguate the speech material is in line with the *cue trading* hypothesis (Streeter, 1978; Beach, 1991) assuming that different prosodic cues contribute jointly to a prosodic effect (that of a boundary marking) and may partially substitute one another such, that in certain contexts one prosodic correlate may be traded in favor of another one. Cue trading phenomena appear in intonation languages, including English, German and Dutch, where for example, post-focal f0 range is frequently compressed, increasing the likelihood of duration and intensity being employed to indicate prominence in the remaining part of the sentence. As a result, a prosodic cue weighting that was primarily identified as being crucial for a particular language, shifted. The phenomena of cue trading is thought to be of particular relevance in manipulated speech as the use of phonetic parameters may follow mechanisms that are different to those of natural stimulus material. Listeners that are faced with a reduced prosodic cue inventory compensate by relying on those cues, that are still available when, for example, assigning a syntactic structure. To summarize, in Experiment 2, the information of the durational properties were sufficient prosodic boundary demarcation cues in perception, though speakers significantly altered their production and used changes of the intonational contour as a salient cue in signaling the sentence meaning.

Boundary tone information

An important acoustic correlate in cueing prosodic boundaries are pitch excursions at the edge of prosodic boundaries (Pierrehumbert, 1980). In contrast to absent pitch information that did not influence parsing substantially, inverting the pitch accent at the position of NP1 (while retaining the durational information) did significantly deteriorate listeners' disambiguation ability. This grammatical (phonological) violation (a dative structure with a rising accent on NP1 followed by a continuation rise (L*H H%) and a genitive structure with a falling accent on NP1 (H*L)) fooled listeners in such a way that they

[3] The results of Experiment 2 do not allow the reverse interpretation that the information of f0 is in principle irrelevant for disambiguating the sentences. Instead, in absence of cues provided by f0, the information of the prefinal lengthening and pause duration provides enough prosodic cues in order to establish the correct sentence interpretation.

accomplished the sentence completion task at chance level (Experiment 1, non-original conditions). In principle, the durational properties of the speech material are, by themselves, informative enough (Experiment 2), so that listeners should have no problems in disambiguating the conditions with inverted boundary tone information. The fact that they did not means that the grammatical violation of the boundary tone information seems to block listeners' recourse to the durational properties of the speech material and enables listeners to properly parse the sentence. In this sense, it becomes evident that interpretative constraints attribute meanings to tones, as suggested by Féry et al. (2009). For the present case, the bitonal rising accent with a followed high boundary tone (L*H H%) of the non-original dative condition, for instance, is interpreted as a non-final accent. In the same way, the bitonal falling accent of the non-original genitive condition (H*L) on NP1 is interpreted as a low boundary tone. In hearing such salient tones on the stressed syllable of NP1, hearers develop clear anticipations of as yet unperceived sentence material, and, therefore, expect more sentence material to integrate into the current prosodic phrase in the non-original dative condition. At the same time, they expect no further sentence material to be included into the non-original genitive condition. This expectation clashes with the actual sentence continuation in both conditions and induces the phonological violation, implying that the sentence-final boundary tone associated with the ι-boundary seems to be linked to semantic/pragmatic meanings, and have been the principal means by which the two meanings of the sentence were differentiated.

Experiments on the perception of German boundary tones evidenced that listeners are highly sensitive regarding the tonal height of an ι-boundary tone. Schneider & Lintfert (2003); Schneider & Möbius (2005) examined the categorical status of the low and high boundary tone in an experiment using sentences that were syntactically ambiguous between statement and question (*Steht alles im Kochbuch* (It's all in the cookbook). The height of the phrase final F0 contour was manipulated on a stimulus continuum ranging from low to high, and listeners had to categorize whether they hear a statement, a question or neither nor. The results found there show that the most important cue for the perception of a statement is the low boundary tone. For the perception of a question, in turn, no such clear results were gained. The results for the perception of the high boundary tone are explained by postulating a third boundary tone called continuation (%) which ..."differs from L% in the height of F0 and from H% only in being non-terminal. A high but non-terminal boundary tone signals that the speaker will continue with his turn, whereas a high and terminal boundary tone (H%) signals that the listener may take the turn" (Schneider & Möbius, 2005). In that sense, the result confirms the perceptual importance of the ι-boundary tone and its interpretative content with respect to upcoming sentence material.

Other studies, investigating the perception of manipulated tones on parsing performance, found a general tendency that misplaced accents induce more difficulty than exaggerated or insufficient ones (Gussenhoven, 1983; Birch & Clifton, 1995; Hruska et al., 2001). Also, it has been shown that listeners are reasonably sensitive to the difficulty of processing incorrect prosodic information as compared to merely infrequent prosodic information (Mietz et al., 2008). In an experiment using ERPs which have been shown to be a valuable tool for the investigation of the time course of prosodic influences on language perception and the temporal dimensions of processing difficulties caused by syntax-prosody-mismatches, Mietz et al. (2008) found that listeners perception of a truly incorrect prosodic contour (syntax-prosody mismatch) induced integration difficulties as reflected by an N400 as well as a reanalysis of a formerly assigned syntactic structure, reflected by a P600. On the other hand, infrequent prosodic information elicited only an N400 similar to words with a low frequency in occurrence (Van Petten & Kutas, 1987), where the N400 is thought to arise from mismatches between expectations on an intonation contour which are built up on-line and the prosodic pattern listeners actually encounter (i.e. a target-actual comparison).

Taken together, only few studies scrutinized the perception of qualitatively different accents and their anticipatory effects, such as Féry & Stoel (2006). However, the obtained results show a consistent pattern that incorrect and misplaced accents are rejected while inappropriate accents are rather tolerated. These general results support the assumption that, for the study at hand, the correct boundary tone information is, again, the crucial piece of information for correctly parsing structures containing an ι-phrase boundary, supporting the observation that f0 is considered as the most important perceptual cue for various prosodic categories, including ι-phrase boundaries (Beckman & Pierrehumbert, 1986a; t' Hart et al., 1990).

Prefinal lengthening vs. pause duration

Prefinal lengthening and pausing are tightly related and have been claimed to contribute to a single percept of pause or boundary (Wagner & Watson, 2010). Perception studies revealed that listeners report hearing pauses even when there are no unfilled pauses in the signal (Martin, 1970b). O'Malley et al. (1973) found evidence that different amounts of pause duration can code different degrees of boundary, a finding that was confirmed in Fant & Kruckenberg (1996) as well as in the production experiments of the present work, where a smaller pause duration correlated with the occurrence of a φ- boundary and a larger pause with that of an ι-boundary. For the current experiments that attested an influence of the durational properties on sentence disambiguation, the intonational contour was removed from the speech. The remaining durational information was of help in signaling the correct sentence condition, and out of the durational information, pause duration between the first and second NP, as opposed to prefinal lengthening on the first NP, appeared to be the prosodic cue representing a stronger prosodic impact on listeners' disambiguation. This is evidence that *pause duration* needs to be detached from *prefinal lengthening*, instead of collapsing both information sources when establishing a specific cue weighting. A potential explanation for the fact that listeners ability to attend to *pause duration* over *prefinal lengthening* when processing ι-phrase boundaries could be found in the different magnitude: speakers significantly altered their use of *prefinal lengthening* and *pause duration* to indicate the meaning of the two syntactic structures, however, the most salient prosodic difference (when neglecting the intonational contour) was certainly the difference in pause duration between NP1 and NP2 (cf. Figure 5.3), which is, in turn, the most prominent prosodic information listeners' rely on when disambiguating the sentence structures. While it seems that the difference of *prefinal lengthening* may not be perceptible, it is used in production. In connection with the segmental phonology, the term 'covert contrast' (Scobbie et al., 1996) has been used to describe situations in which a prosodic contrast is produced between two categories, but because the realizations fall within one perceptual category, the contrast is not recognized. This may the case for the described difference of *prefinal lengthening*. Analyzing production and perception data so tightly can reveal these imperceptible differences and can tell us something about the cognitive nature of the transition from the production level to the perception level.

Generally, there is a paucity of work on how listeners weight durational cues to prosodic boundaries. However, the small amount of cross-linguistic research that has been done, suggests a diversity. English speaking adults, for example, seem to rely more strongly on *prefinal lengthening* than on pause in detecting clausal boundaries (Aasland & Baum, 2003). Likewise Russian adults do not weight pause very heavily (Volskaya, 2003). In contrast to this, Dutch (Sanderman & Collier, 1997) and Swedish speakers (Horne et al., 1995; House et al., 1998) do rely heavily on pause duration for prosodic boundary judgments. Further evidence from Chinese listeners, in turn, suggest that they weight pitch changes over durational cues when perceiving prosodic boundaries (Xinting, 2011). In that context, German seems to capture an intermediate level in that the use of prosodic correlates is boundary size dependent, with *prefinal lengthening* used for detecting φ-boundaries and *pause duration* used for identifying ι-phrase boundaries under the absence of cues provided by f0. Importantly, the role of a particular acoustic cue is not universal, since certain prosodic phenomena are realized differently across languages, see Hirst & Di Cristo (1998). Such cross-linguistic differences in perceptual saliency indicate that language exposure modulates listeners focus to cues that are particularly relevant in the native language.

To summarize, the experiments were constructed to test the hypothesis that listeners prefer to perform a prosodic cue weighting depending on the strength of the boundary, instead of ranking their cues invariable in different boundary environments. The sum of all experiments speak against the strategy that listeners need the holistic set of prosodic parameters for successful disambiguation. Rather, it appears that the cue use is dependent on the size of the prosodic boundary.

The next section presents how the generalization that prosodic correlates of φ- and ι-boundaries are differently used in perception, and derives a grammar for both boundary environments. The mechanism for the derivation of this generalization is taken from the inventory of constraint formulation within Optimality Theory (OT).

6
Theoretical evaluation of the findings

The following section attempts to account for the observed prosodic cue weighting on φ- and ι-boundaries by implementing the results in an Optimality-Theoretic (OT) model. Solely making use of constraints derived from the results of the perception experiments presented in the previous chapters, the OT-models are capable of capturing the processing data by advocating one relevant cue per boundary size that determines the mapping of a phonetic input onto a phonological structure. The models predict that in the face of a φ-boundary, lengthening constraints determine the respective mapping, while the involvement of an ι-phrase boundary calls for the boundary tone itself to determine syntactic ambiguity resolution.

6.1 Standard Optimality Theory

Within Optimality Theory, an input structure is mapped onto an output structure via two functions (Prince & Smolensky, 1993, 2004). The first of these two functions is a generative component (GEN) that takes an input drawn from the lexicon, and creates a set of alternative realizations of the input as output candidates. These output candidates are then subjected to the evaluative function (EVAL) that compares the generated candidate set among each other and selects the best (grammatical) candidate as the output. In order to select the best candidate, EVAL consults an universal set of violable grammatical constraints which conceptualize linguistic principles and requirements. Two kinds of grammatical constraints can be distinguished in OT: markedness constraints and faithfulness constraints. The term 'markedness and faithfulness' as applied to constraints was coined in Prince & Smolensky (1993). Markedness constraints regulate each output form with respect to a given linguistic principle. For example, NUC and NOCODA are markedness constraints, requiring unmarked syllable structure in the generated output candidates. According to NUC, syllables must have a nucleus, and according to NOCODA, they should not have codas, favoring structural configurations in the output representation over others (e.g. syllables without a coda over syllables with a coda, see Prince & Smolensky (1993) and McCarthy & Prince (1993) for detailed accounts of the syllable structure in OT). In contrast, faithfulness constraints require a close relationship of the output representation to its corresponding input. For example, MAX demands no deletion of segments, and DEP forbids epenthesis. Such linguistic principles are formulated as constraints which are universal and present in the grammar of all languages. These constraints impose conflicting requirements on the output candidates. Output candidates, for example, that deviate with respect to MAX and DEP from a specified input form are rejected and eliminated as optimal candidates by these faithfulness constraints. Obviously, no output candidate can possibly obey all constraints. Instead, the candidate set is evaluated against a universal constraint set obeying a language-specific ranking. The output candidate encountering the fewest violations of the highest ranked constraints is selected as the optimal candidate.

6.2 OT on phrase structure formation

Among the endeavors first approached in OT were accounts on mapping syntactic XPs to prosodic structure which since then serve as a starting point example for the application of standard OT in syntax-prosody relations (Selkirk, 1995), and which are also used in the current section as an example how to derive phrasal patterns from syntactic inputs. The cross-linguistic observation that syntactic structure can be mapped onto prosodic structure led Selkirk (1986) to suggest a universal theory of phrasing based on alignment of φ-phrases with syntactic XPs according to certain interface constraints. Later McCarthy and Prince (1993) generalized edge alignment in form of generalized alignment constraints, shown in (18). Applying the same constraint formalism, Selkirk (1995) introduced edge-alignment constraints, that require alignment of the left or right edges of XPs with those of phonological phrases (φ-phrases) (19). Her account constitutes an influential analysis in OT syntax-prosody mappings, and a majority of succeeding analyses within OT refers back to her works.

(18) Generalized Alignment
Where *Cat1, Cat2* are prosodic, morphological, or syntactic categories and *Edge1, Edge2* $\in$\{Right, Left\}:
$\text{Align}(Cat1, Edge1; Cat2, Edge2) \Leftrightarrow$
"For each *Cat1* there is a *Cat2* such that *Edge1* of *Cat1* and *Edge2* of *Cat2* coincide."

(19) a. ALIGN-XP,R: ALIGN(XP, R; P, R)
"For each XP there is a P such that the right edge of XP coincides with the right edge of P."

b. ALIGN-XP,L: ALIGN(XP, L; P, L)
"For each XP there is a P such that the left edge of XP coincides with the left edge of P."

Not all existing phrasal patterns can be entirely derived by applying edge-alignment constraints (Truckenbrodt, 1995). It is argued that another type of interface constraint is relevant. Truckenbrodt (1995, 1999) proposed a constraint on the syntax-prosody interface, the WRAP constraint, which requires material of a syntactic constituent type *a* be included or wrapped within a prosodic structure constituent type *b*. The character of the WRAP constraint is cohesional: a syntactic constituent whose content is represented into separate prosodic constituents causes a violation of the constraint. Truckenbrodt (1999) formulated the constraint WRAP-XP as in (20).

(20) WRAP-XP
"Each XP is contained in a phonological phrase."

The syntax-prosody-mapping in terms of ALIGNMENT and WRAP constraints can be modeled in the framework of Optimality-Theory. Consider the hypothetical configuration of a verb followed by two noun phrases as shown in Tableau (21).

/ [[Verb] [Noun]NP [Noun]NP]VP /	ALIGN-XP,R	WRAP XP
a. (V NP)ϕ (NP)ϕ		*
b. (V NP NP)ϕ	*	

Figure 6.1: Tableau 21 (unranked constraint formation of Align-XP, R and Wrap XP)

Tableau (21) depicts that WRAP-XP may be in conflict with certain boundaries required by ALIGN-XP,R. In certain structural configurations, the satisfaction of one constraint results in the violation of the other. We can see this in Tableau (21), where two output candidates are given. Candidate a. contains two prosodic phrases (φ–phrases) with a phrasal break between the complements of the verbal phrase (VP). In candidate b., the entire VP is contained in one φ-phrase. Candidate a. satisfies ALIGN-XP, R by coinciding edges of φ–phrases with all right edges of syntactic phrases appearing in the input. However, candidate a. does not meet the requirements demanded by WRAP-XP since the entire VP is not wrapped into one single φ–phrase, and is therefore marked with a '*' in the corresponding cell, following the OT standard conventions for signaling a violation. Candidate b. triggers one violation of ALIGN-XP,R since the right edge of the first NP in the input does not coincide with a φ-phrase in the output. Though, candidate b. satisfy WRAP-XP since there is a single φ–phrase containing the entire VP. For this hypothetical configuration it is assumed that ALIGN-XP,R and WRAP-XP are unranked in the grammar, as indicated by the dotted line between the constraint columns.

Truckenbrodt (1995) observed that languages like ChiMwiini or Chaga show Align-XP effects, and proposed that ALIGN-XP is higher ranked than WRAP-XP: ALIGN-XP » WRAP-XP, which is commonly indicated in the OT-Tableau by a solid line between the constraint column. In these languages, this constraint ranking is responsible for the violation of WRAP-XP, as well as for selecting candidate a. out of our hypothetical candidate set in Tableau 18 as the optimal candidate. For languages like Chizigula, in turn, in which the effects of ALIGN-XP appear to be different, Truckenbrodt (1995) proposed that the ranking of the two constraints is reversed: WRAP-XP dominates ALIGN-XP in the constraint hierarchy: Wrap-XP » Align-XP. Consequently, candidate a. fatally violates the higher ranked constraint WRAP-XP, while candidate b. meets the requirements of WRAP-XP and turns out to be the optimal candidate. This illustrates an example of the sort of typology made available in Optimality-Theory to derive prosodic phrase structures, assuming an universal constraint repertoire. Possible cross-linguistic differences are just those that would derive from differences in constraint ranking.

With respect to listeners' perception of the syntactic structures to be evaluated, i.e. genitive-dative-ambiguity, it is assumed that first, a phonological structure is generated on the basis of pure phonetic information, this process will be referred to as phonetics-phonology mapping. The output of the phonetics-phonology mapping is a phonological structure serving as input representation for the following syntax-prosody mapping. We will first describe the phonetics-phonology mapping in the following section, before the corresponding syntax-prosody mapping for the ambiguity under scrutiny is considered in a model of phonological perception (Boersma & Hamann, 2009; Féry et al., 2009).

6.3 Phonetics-Phonology mapping

The production experiments presented in chapter 4 and 5 evidenced a significant prosodic difference in the vicinity of the first and second NP, and at the position of the prosodic phrase boundary itself. It was furthermore shown that these differences in prosodic rendition guide listeners towards the correct structure of the two possible available computations in perception, and, more specifically, that the use of such prosodic cues varies with the size of the boundary involved in the particular structure. On the basis of these findings in perception and of existing models in phonological perception in Optimality Theory (Flemming, 1995; Boersma, 2009; Boersma & Hamann, 2009; Féry et al., 2009), a model of the perception of boundary related phonetic cues is proposed in the following section. The results are relevant for a theory of perception of auditory inputs as they argue and account for the importance of different boundary cues related to different boundary sizes, and furthermore provide specific evidence for how one relevant phonetic cue guides the mapping of a phonetic input onto a phonological structure. The following section introduces the formal aspects of the phonetics-phonology mapping as assumed in Boersma & Hamann (2009); Féry et al. (2009) among others, and indicates the processing levels in perception assumed for the sentence material investigated in this dissertation.

In general, the perceptual process can be described by referring to (at least) three levels: the lowest level which Boersma (2009); Boersma & Hamann (2009) call the phonetic level, the intermediate level which is referred to as the surface form and the highest level called underlying form, as shown in Figure 6.2. Figure 6.2 depicts how the perceptual process works. It is assumed that listeners' comprehension starts out from an auditory phonetic signal and terminates at an underlying phonological representation. This phonetics-phonology mapping, also referred to as *prelexical perception* or *phonetic parsing*, is guided by grammatical cue constraints that evaluate the relation between the phonetic form and the phonological surface form. Cue constraints were already assumed in models by Boersma (1998, 2000); Escudero & Boersma (2003); Escudero (2005), where a specific cue ranking expresses the details of a language specific mapping from auditory cues to phonological elements. For the present case of the phonetics-phonology mapping, the phonetic level corresponds to the phonetic signal the listener is exposed to, it is the raw F0 contour and/or the phonetic length of the speech signal. The surface form corresponds to an abstract phonological representation, (in other cases, the surface form represents pitch accents, boundary tones or lexical tones, for example), and the underlying form represents the meaning of the linguistic unit; see also Féry et al. (2009) for the assumption of three levels in tonal perception and in particular for arguments in favor of the surface form. The suggestion of three levels in perception is grounded on the idea that processing proceeds in two separate steps (McQueen & Cutler, 1997). The first step also referred to as *perception per se* (Féry et al., 2009) or *identification* (Boersma, 2009) converts the auditory phonetic signal into a phonological abstract surface form. The second step which can be equated with *recognition* or *phonological comprehension* interprets the abstract surface form and assigns a meaning (genitive or dative in the present example).

To summarize, it is assumed that the phonetics-phonology interface is located in a connection between auditory form and phonological surface form. Both levels of representations are linked via cue constraints (Boersma, 1998) that are seen to govern listeners *prelexical perception*, among other cues such as faithfulness constraints, structural constraints and lexical constraints relevant for transforming the auditory signal over the surface form into the final underlying form (Boersma, 2009). For a detailed description of each constraint type and its application in the process of comprehension, the reader is referred to Boersma (2009). We concentrate here on cue constraints only as they are crucial for the mapping of auditory phonetic to surface form. One example how cue constraints guide the phonetics-phonology mapping is the vowel duration in English. In English as in many other languages there is an overall tendency to lengthen a vowel preceding a voiced obstruent (House & Fairbanks, 1953; Peterson & Lehiste, 1960) and perception studies showed that English listeners use the vowel duration as a cue to predict the voicing of the following obstruent (Denes, 1955; Hogan & Rozsypal, 1980). Similarly, it has been shown for German that tonal cues such as a rising or falling F0 contour regulate the mapping of a phonetic input representation onto a phonological surface structure. In a second step, this phonological structure is assigned a semantic-pragmatic meaning such as *topic* or *focus*, as in the example illustrated in Figure 6.2. This meaning information can be employed by listeners to setup expectations of the discourse status of upcoming referents (Féry et al., 2009). Figure 6.2 below illustrates a model of tone perception, as suggested by Féry et al. (2009). A listener who follows Figure 6.2 starts out with a phonetic form (here a rising F0 contour) and can sequentially process the phonological structure (a rising pitch accent or tonal morpheme (Liberman, 1975; Pierrehumbert, 1980; Ladd, 1996)) with the final aim of arriving at a semantic, meaningful interpretation (Topic). Note that the process of production could be implemented as the reverse application of rules until the speaker ends up with a phonetic form (Liberman & Douglas, 2000), as illustrated on the right stack in Figure 6.2.

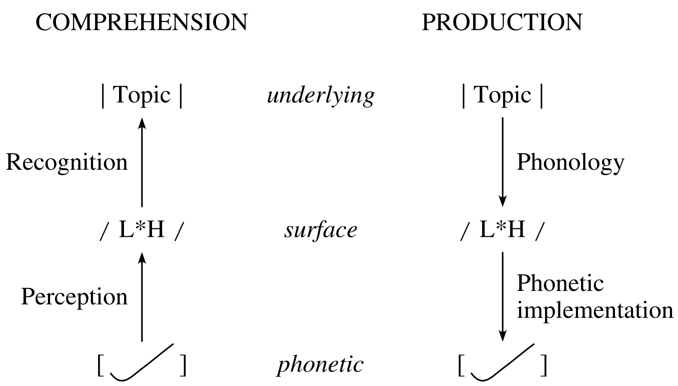

Figure 6.2: Model of tone perception, from Féry et al. (2009: 256)

Crucial for the present model is the transformation of the phonetic representation into an abstract phonological object/category, i.e. the process of *perception*. This operation measures what the listener considers as the speakers most likely intended language specific phonological surface structure, while incrementally parsing the phonetic representation. As illustrated above, *perception* operates via consulting language specific constraints. The OT inherent principle of violable constraints suggests that the degree of obedience to the grammatical constraints can be assessed on complete sentences as well as on sentence fragments or any uncompleted string of representation. This property makes OT as an ideal system to model linguistic competence, in particular when dealing with unfinished locally ambiguous sequences (choice points), as in the present case. The choice point in the present sentence material is located at the position preceding the ambiguous NP2, i.e. the location of the possible occurrence of a φ-boundary, an ι-phrase boundary or no phrase boundary at all. In the perception experiments presented in the previous chapters, we identified one relevant phonetic cue in dependence of each prosodic boundary. As discussed and summarized at the end of section 4 and 5, *final lengthening* as boundary cue guides listeners' perception of a φ-boundary, and the *boundary tone* turned out to be a necessary boundary parameter for listeners' perception of an ι-boundary. The two prosodic correlates seem to be the most important in the distinction between the two boundary types, other occurring cues are subordinate and of minor relevance. This pattern of cue weighting for the particular prosodic boundary is shown in (22) and (23), illustrating the perceptual weight/influence each prosodic cue constitutes in ambiguity resolution, as conveyed from the statistical analysis of the perception experiments presented in the previous chapters.

(22) φ-boundary:
Final-Length > Duration-Pause > Boundary-Tone (f0)

(23) ι-boundary:
Boundary-Tone > Duration-Pause > Final-Length

Thus, two boundary correlates appear to be necessary[1] and enter our model as the most important cues responsible for the phonetics-phonology mapping. In OT, such cues are formalized as cue constraints. For the evaluation of φ-boundaries, the relevant constraints are listed in (24), (25) and (26). The first two constraints can be understood as markedness constraints that regulate the output forms with respect to a given linguistic principle. (24) and (25) refer to the fact that the identification of a φ-boundary following NP1 requires segmental material that is prefinally lengthened. Consequently, segmental material that is not lengthened at the position of NP1 does not imply finality in terms of a φ-boundary. Therefore, a family of Length constraints militates in favor of perceiving phonetic inputs in a categorical manner.[2] Thus, Length-Fin [> = 500ms] requires the listener to treat an auditory input with a duration of 500ms and longer as member of a category that implies finality (FIN) indicated as a φ-boundary after NP1 (24). Length-Fin [> = 500ms] is a constraint referring to a minimum demanding that auditory inputs with a minimum of 500 ms are interpreted as indicating finality (FIN) of a φ-phrase. Similarly, Length-Con [< = 400ms] requires the listener to treat an auditory input with a duration of 400ms or shorter as member of a category that implies continuation (CON) indicated by the absence of a φ-boundary after NP1 (25). Thus, Length-Con [< = 400ms] refers to a maximum such that auditory inputs with a maximum of 400 ms are interpreted as indicating continuity (CON) of a φ-phrase. Upper and lower limit duration values for the constraints Length-Fin [>=500ms] and Length-Con [< = 400ms] resulted from the original duration values on NP1 measured in the production experiment, as well as from the recordings that served as auditory inputs to the perception experiments.[3] The third constraint in (26) is a faithfulness constraint requiring a close relationship of the output representation to its corresponding input. Specifically, (26) militates against output forms that are not faithful to the lengthening values of the input form.

[1] A prosodic cue is regarded as *necessary* in the present work when its manipulated phonetic value lead to such a deteriorated disambiguation performance that the two syntactic structures of the ambiguities examined in chapter 4 and 5 can not be ascertained anymore.

[2] See Boersma et al. (2003) who assume a family of Perceive constraints to model the mapping from auditory inputs to phonetic categories in OT.

[3] Exact duration values for each NP1 value are given in Appendix C.

(24) LENGTH-FIN [> = 500MS]
"Auditory inputs with 500ms or longer are treated as members of a category that implies finality (FIN) indicated by a φ-boundary."

(25) LENGTH-CON [< = 400MS]
"Auditory inputs with 400ms or shorter are treated as members of a category that implies continuation (CON) indicated by the absence of a φ-boundary".

(26) IDENT-IO (LENGTH)
"Input and output correspondents have the same specification for length".

The prosodic constraints introduced in (24), (25), and (26) are the core components of the model that are said to regulate the phonetics-phonology mapping and are needed for the Optimality-Theoretic model of perception. Optimality-Theoretic Tableaus in (28), (29) and (30) illustrate the grammar responsible for the mapping procedure of a phonetic input representation onto a phonological surface structure, i.e. the process described as *perception* in Figure 6.1. Input of the following models is a phonetic representation/auditory signal with varying lengthening values of NP1. The output consists of both a phonological phrase structure featuring either or a φ-boundary after NP1, and varying lengthening values of NP1. Three constraints introduced in (24)-(26) determine the phonetics-phonology mapping; as illustrated in (28)-(30), the constraints are not ranked here as they do not interfere with each other. Following common notations of the representations, square brackets are used for phonetic forms, and slashes for phonological forms (Boersma, 2009). The model will be applied to NP1, i.e. the position preceding the choice point of the case ambiguity in (27), where (27a) features a φ-boundary after NP1 and (27b) does not. All together, three different input representations will be evaluated in (28)-(30) that vary in their duration of NP1. In what follows, an OT model is outlined that makes falsifiable predictions for phonological phrase structure generation at the position of NP1.

(27a) dative-reading
Neulich hat (der Gärtner$_{NP1}$)φ (der Reiterin$_{NP2}$)φ den Baum gezeigt,...
Recently the gardener showed the horsewoman a tree, ...

(27b) genitive-reading
Neulich hat (der Gärtner$_{NP1}$ (der Reiterin$_{NP2}$)φ)φ den Baum gefällt,...
Recently the gardener of the horsewoman chopped a tree, ...

Perception of φ-boundaries

Tableau 28 illustrates the first of two successive NPs as input with a duration of 600ms length. Three possible output candidates with varying length for NP1 are represented, either featuring a φ-boundary or not, summing up to six possible output representations. LENGTH-FIN [> = 500MS] claims a φ-boundary after each NP whose durations is 500ms or longer. Accordingly, as for candidate c., the appearance of an NP1 duration of 600ms followed by a φ-boundary suspends the violation of LENGTH-FIN [> = 500MS]. All

	[NP1 = 600 ms]	LENGTH-FIN [> = 500 ms]	LENGTH-CON [< = 400 ms]	IDENT-IO (LENGTH)
	a. / (NP1)ϕ 350ms /	*	*	*
	b. / (NP1)ϕ 450ms /	*		*
☞	c. / (NP1)ϕ 600ms /			
	d. / (NP1 350ms /			*
	e. / (NP1 450ms /		*	*
	f. / (NP1 600ms /	*	*	

Figure 6.3: Tableau 28 [der GärtnerNP1 der ReiterinNP2]

candidate structures with a NP1 duration shorter than 500ms but followed by a *φ*-boundary (candidate a. and b.) incur a violation of Length-Fin [> = 500ms]. Similarly, those candidates with a duration of 500ms and longer featuring no *φ*-boundary (candidate f.) violate Length-Fin [> = 500ms] as well and are out of bound. Recall, Length-Con [< = 400ms] demands no *φ*-boundary after each NP whose duration is 400ms or shorter. Correspondingly, candidate a. does not meet the requirements of Length-Con [< = 400ms] as it shows a *φ*-boundary after NP1 in spite of a duration of 350ms. Similar, candidate e. and f. violate Length-Con [< = 400ms] as they feature no *φ*-boundary although their NP1 duration exceeds 400ms. Consequently, they are ruled out as possible candidates as well. A duration of NP1 disparate to the input duration of NP1 is prohibited by Ident-Io (Length). Accordingly, candidate a., b., d., and e. violate Ident-Io (Length). In what follows, candidate c. represents the optimal candidate with a duration of 600ms for NP1 followed by a *φ*-boundary.

Turning now to the phonetics-phonology mapping for an input duration of 350ms length (Tableau 29). Those candidates with an NP1 duration shorter than 500ms featuring a *φ*-boundary (candidate a. and b.) and likewise those candidates with a NP1 duration of 500ms or longer (candidate f.) incur a violation of Length-Fin [> = 500ms]. Therefore, candidate a., b. and f. are ruled as optimal output candidates. Candidate a. furthermore does not meet the requirements of Length-Con [< = 400ms] by showing a *φ*-boundary after NP1 in spite of a duration of 350ms. Just as illustrated in Tableau (25), candidate e. and f. violate Length-Con [< = 400ms] as well, as these candidates appear without a *φ*-boundary albeit their NP1 duration exceeds 400ms. All candidates that are not equal in duration to the input representation (candidate b., c., e. and f.) fail due to a violation of Ident-Io (Length). As a result, for an input duration of 350 ms for NP1, candidate d. turns out to be the winner.

[NP1 = 350 ms]	LENGTH-FIN [> = 500 ms]	LENGTH-CON [< = 400 ms]	IDENT-IO (LENGTH)
a. / (NP1)ϕ 350ms /	*	*	
b. / (NP1)ϕ 450ms /	*		*
c. / (NP1)ϕ 600ms /			*
☞ d. / (NP1 350ms /			
e. / (NP1 450ms /		*	*
f. / (NP1 600ms /	*	*	*

Figure 6.4: Tableau 29 [der GärtnerNP1 der ReiterinNP2]

Now, consider the case with an input duration of 450 ms for NP1 (cf. Tableau 30). In contrast to the previous Tableaus (28) and (29) where one of the output candidates satisfies all constraints and thus can be selected as the winner, in (30) output candidates with a duration identical to the input duration (potential winning candidates) certainly satisfy IDENT-IO (LENGTH) (candidate b. and e.), they however, violate either LENGTH-FIN [> = 500MS] when featuring a *φ*-boundary (candidate b.) or they incur a violation of LENGTH-CON [< = 400MS] when appearing without a *φ*-boundary (candidate e.). Therefore, none of the output candidates satisfy all three constraints as they either violate one of the Length-Constraints or they founder on IDENT-IO (LENGTH) for reasons described above. Instead, we have four optimal candidates with one violation each. We have seen in Tableau (28) and (29) above that LENGTH-CON [< = 400MS] does not contribute in determining the optimal candidate; the work is done by restrictions expressed through LENGTH-FIN [> = 500MS] and IDENT-IO (LENGTH) only. However, at this point it becomes evident why we need the constraintLENGTH-CON [< = 400MS] at all. This constraint is responsible for the failure of identifying an optimal candidate when the input duration of NP1 floats between the clear-cut boundaries of a 'dative NP1 duration' and that of a 'genitive NP1 duration'. This is a clear case where a phonetic input signal cannot be unambiguously transformed into one type of a phonological object (either that of candidate b. or e.) and therefore, also the process of comprehension (referred to as 'recognition' in Figure 6.1 above) which interprets these phonological objects as meaningful concepts is not a clear-cut one. In other words, an input representation of 450ms for NP1 constitutes an ambiguous case, and as a parsing consequence, listeners' disambiguation ability deteriorates. At that point we can predict that listeners will not disambiguate the ambiguous sentence fragment at chance level but instead resolve the ambiguity primarily towards the dative structure (see chapter 4, Experiment 3, Figure 4.8). A parsing preference for the dative reading based on syntactic properties had already been established in chapter 3.

[NP1 = 450 ms]	LENGTH-FIN [> = 500 ms]	LENGTH-CON [< = 400 ms]	IDENT-IO (LENGTH)
a. / (NP1)ϕ 350ms /	*	*	*
☞ b. / (NP1)ϕ 450ms /	*		
☞ c. / (NP1)ϕ 600ms /			*
☞ d. / (NP1 350ms /			*
☞ e. / (NP1 450ms /		*	
f. / (NP1 600ms /	*	*	*

Figure 6.5: Tableau 30 [der GärtnerNP1 der ReiterinNP2]

Perception of ι-boundaries

The phonetics-phonology mapping at the level of the ι-phrase, in turn, is guided by tonal properties at the end of NP1. This claim is based on perception results presented in the previous chapter indicating that the perception of the tonal contour at the position of the choice point is the most crucial cue listeners draw on when disambiguating the ambiguous sequences (cf. chapter 5, Experiment 1, Figure 5.5). The tonal contour in the dative-reading at the position of NP1 was phonetically grounded as a bitonal falling accent (H*L) followed by a low (L%) boundary tone, while it was a bitonal rising accent (L*H) in the genitive reading. A manipulation of the tonal contour at the position of NP1 resulted in two further manipulated conditions exhibiting for the dative reading a bitonal rising accent (L*H) followed by a high (H%) boundary tone, and for the genitive reading a bitonal falling accent (H*L), see Table 6.1.

Condition	NP1
(a) original dative	H*L L%
(b) original genitive	L*H
(c) manipulated dative	L*H H%
(d) manipulated genitive	H*L

Table 6.1: Tonal contour at the position of NP1.

In German, as in many other languages, the contrast between the low (L%) and the high (H%) boundary tone corresponds to a contrast in sentence mode and represents two prosodic categories (Schneider, 2012). The low boundary tone is interpreted as a statement while the high boundary tone is interpreted as a question, signaling *finality* in the first case and *turn taking* or *continuation* in the latter one. We extend the argument that, similar to the categorical status of the low (L%) and the high (H%) boundary tone in German and its respective interpretation, German bitonal pitch accents convey a similar categoriality in meaning. That is, a bitonal falling accent (H*L) of a phrase get an interpretation which expresses *finality*, and a bitonal rising accent (L*H) of a phrase is interpreted as *non-final* or as *continuous*. This is not only evidenced by the perception experiments in the previous chapter 5 but relies on well-established facts on German intonation (Büring, 1997; Grabe, 1998; Grice et al., 2005; Féry, 1993) summarized in Féry et al. (2009: 244), namely that "(i) a declarative sentence has a global falling contour, (ii) a bitonal falling accent is the last one in a declarative Intonation Phrase, and (iii) a bitonal rising tone is not the last in a declarative sentence", among other tonal phenomena.

Accordingly, we argue that a falling accent (H*L) followed either with or without a low boundary tone (cf. condition (a) vs. (d) in Table 6.1 above) falls into the same perceptive category, namely that of signaling *finality*. Similarly, the rising accent (L*H) followed either with or without a high boundary tone (cf. condition (b) vs. (c) in Table 6.1 above) fall into the same perceptive category as well, namely that of signaling *non-finality* or *continuation*. Falling tonal contours are indicated with an ι-boundary at its right edge, whereas rising tonal contours expresses phrasal continuity and are therefore not associated with an ι-boundary at its right edge.[4] To model the perceptual behavior on ι-boundaries, two markedness constraints and one faithfulness constraint are needed (31)-(33), that operate at the phonetics-phonology interface. The first two markedness constraints (31) and (32) refer to the observation that the presence of an ι-boundary depends on the preceding tonal contour of the lexical element. Specifically, a falling contour formalized as [HL] is associated with a low ι-boundary at its right edge. In other words, a [HL] contour that is not followed by a right ι-boundary is prohibited (31). Similarly, a rising contour described as [LH] is associated with phrasal continuity, i.e. with the absence of an ι-boundary at its right edge. Consequently, a [LH] contour that is followed by a right ι-boundary is prohibited (32). Similar to (26), (33) is a faithfulness constraint operating at the phonetics-phonology interface, and requiring a close relationship of the output representation to its corresponding input in that it militates against output forms that are not faithful to the tonal values of the input form.

(31) *ι(HL
"[HL] contours not followed by a right ι-boundary are prohibited."

(32) *LH)ι
"[LH] contours followed by a right ι-boundary are prohibited."

(33) IDENT-IO (TONE)
"Input and output correspondents have the same specification for tone".

In order to model the perception of ι-boundaries, three constraints introduced in (31), (32) and (33) are necessary to regulate the phonetics-phonology mapping. Two Optimality-Theoretic Tableaus with a varying tonal input representation depict the grammar responsible for the mapping procedure. Tableau (35) represents a falling contour on NP1 as input representation, indexed as L, and Tableau (36) shows the case of a rising contour on NP1 as input, indexed as H. The generated output candidates represent abstract phonological categories at the surface level, featuring either a falling or a rising contour on NP1, and which are once associated with an ι-boundary and once without an ι-boundary, adding up to a total

[4] Cross-linguistically it has been observed that a rising contour at a boundary signals continuity. This sound-meaning relation between a phrase final rising pitch contour and continuation is generally referred to as *continuation rise* (Cruttenden (1986) for British English, Pierrehumbert (1981) for American English, t' Hart et al. (1990) for Dutch, von Essen (1956) for German, among others). For West Germanic languages, there are two grammaticalised forms of the continuation rise: H% (a high boundary tone) and a H (high movement without a boundary tone), both forms can in turn be preceded by different pitch accents, including H*, H*L, L*H and L* (Chen, 2007). As a result, there is more than one pitch contour that can be used to signal continuation.

of four output representations. The presented model will be again applied to NP1, i.e. the position preceding the choice point of the case ambiguity presented in (34). (34a) features an ι-boundary after NP1, while in (34b) no boundary is postulated at the same position.

(34a) dative-reading
Ständig isst (der Kunde$_{NP1}$)ι. (Der Händlerin$_{NP2}$)φ vergeht das Lachen.
The customer eats constantly. The merchant does not like it.

(34b) genitive-reading
Ständig isst (der Kunde$_{NP1}$ (der Händlerin$_{NP2}$)φ)φ im Obstladen, ...
The customer of the merchant eats constantly in the fruit shop, ...

	[NP1 = HL]	*ι(HL	*LH)ι	IDENT-IO (Tone)
	a. / (NP1 LH /			*
	b. / (NP1 HL /	*		
☞	c. / (NP1) HLι /			
	d. / (NP1) LHι /		*	*

Figure 6.6: Tableau 35 [der KundeNP1: HL at NP1]

Tableau 35 illustrates the phonetics-phonology mapping of a falling contour (HL) on NP1. Four possible output candidates are represented that vary in their tonal realization (high versus low) and whether the NP is coincided with or without an ι-boundary. In (35) candidate c. turns out to be the winner as it obeys all constraints. Candidate a. and candidate d. violate IDENT-IO (TONE) once, since their tonal representations are not identical to the tonal features of the input representation. Candidate d. furthermore incurs a violation of *LH)ι as NP1 is coincided with an ι-boundary prohibited by the respective constraint. Though candidate b. appears to be tonally identical to the input representation, it is ruled out by *ι(HL since it is not featuring an ι-phrase boundary as required. This way, Tableau (35) illustrates how a falling contour on NP1 allows us to formulate prediction about the phrasal structure. A falling contour is optimal only when it is associated with an ι-boundary (candidate c.), but not without a boundary (candidate b.). It is in any case not compatible with a rising contour, no matter whether the rise on NP1 is followed by an ι-boundary (candidate d.) or without a boundary (candidate a.).

[NP1 = LH]	*ι(HL	*LH)ι	IDENT-IO (Tone)
☞ a. / (NP1 LH /			
b. / (NP1 HL /	*		*
c. / (NP1) HLι /			*
d. / (NP1) LHι /		*	

Figure 6.7: Tableau 36 [der KundeNP1: LH at NP1]

In (36) the input representation on NP1 is a rise, as indexed with the subscripted character H. (36) illustrates that a rising contour on NP1 is compatible only without a boundary (candidate a.). Candidate d. featuring a rise as well is however eliminated due to a violation of *LH)ι requiring a rise to be denoted without a boundary. Candidate b. and c. do not meet the requirements of Ident-Io(Tone) and are eliminated as well.

Generally, Tableau (35) and (36) illustrate that a falling contour is associated with an ι-boundary, whereas a rising contour is not associated with an ι-boundary. Note, that these results reflect the specific mapping of a phonetic input signal onto an abstract phonological surface structure for the present case of a syntactic ambiguity and predicts how listeners perceive the phonetic properties of the auditory input signal in order to assign a phonological structure, accordingly. This phonological structure serves an input representation to the syntax-prosody mapping described below. The result of the syntax-prosody mapping in turn is a prosodic phrase structure, corresponding either to that of the genitive condition or that that of the dative-condition.

6.4 Syntax-prosody mapping

This section first motivates the interface-constraints necessary for the the syntax-prosody mapping. Before the relevant constraints and their requirements will be introduced, the syntax of the structures to be evaluated will be presented.

As for the present Genitive/Dative-ambiguity, representing a locally ambiguity composed of a sequence of two successive NPs, the parser has the option of interpreting the second NP either as the complement of the verb (dative object-interpretation), or as a modifier of the first NP (genitive-interpretation). The corresponding syntactic structures of each condition were introduced in chapter 3, section 3.2.1 and are repeated here in Figure 6.2 in order to propose the relevant constraints and its ranking necessary for the derivation of the respective prosodic phrase structure.

The ambiguity at hand contains three NPs, from which the ambiguous second NP *der Reiterin* could either be interpreted as an argument of a ditransitive verb (37), or as a possessive modifier of the first NP, resulting in a complex possessive nominal structure involving a transitive verb (38). Both sentences are initially ambiguous until they are disambiguated by the transitivity information on the verb.

(37) dative structure:
Neulich hat der Gärtner der Reiterin den Baum gezeigt, der morsch war.
Recently the gardener showed the horsewoman a tree, that was decaying.

(38) genitive structure:
Neulich hat der Gärtner der Reiterin den Baum gefällt, der morsch war.
Recently the gardener of the horsewoman chopped a tree, that was decaying.

In (37) the ambiguous NP is analyzed as one of three arguments to the following ditransitive verb. Figure 6.8 on the left illustrates the corresponding syntactic structure of the verbal phrase starting with NP1 up to and including the verb. The verb takes the NP *den Baum* as its complement which gets assigned Accusative case. The dative argument *der Reiterin* occurs as a specifier of VP. The subject NP *der Gärtner* sits in the specifier of little vP. Little v takes VP as its complement. Therefore, NP2 is a full constituent of the core VP since it is selected as an argument by V0.

The alternative interpretation of (37) is shown in (38) and has the syntactic properties pictured in Figure 6.8 on the right. In this case, the ambiguous NP2 is not an argument of the verb. Instead, this NP is embedded within the NP headed by the noun *Gärtner*. The embedded NP serves as a modifier of *Gärtner*. Contrary to arguments, modifiers occur outside the VP and hence are not analyzed as arguments of V0. Recall that for reasons of comparability, we used NP as a label for a noun (N) as well as for a proper noun phrase (NP). We continue to use the same labeling in the following analyses.

Comparing the two syntactic representations in Figure 6.8 by focussing on the region up to and including the disambiguating verb information, the crucial difference outlined above is evident at the position of NP2. This syntactic difference is particularly relevant for the prosodic phrase structure formation presented in the following section.

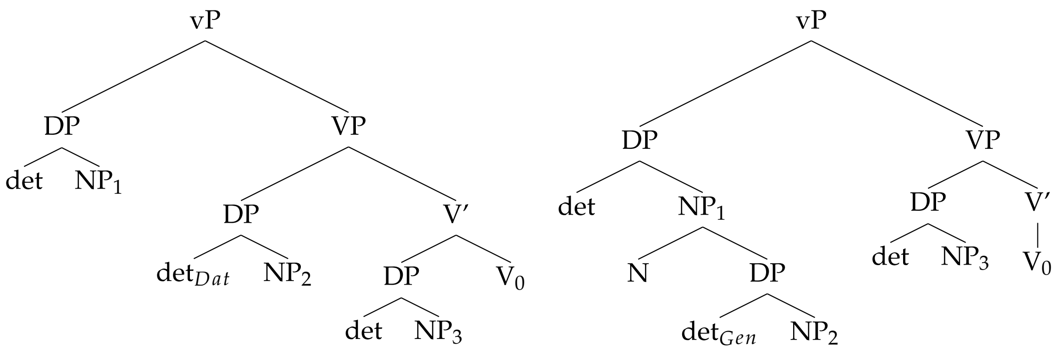

Figure 6.8: Structural representation of the dative-reading, on the left (Chomsky, 1993, 1995, 2000). Example sentence of the dative-reading: *Neulich hat der Gärtner der Reiterin den Baum gezeigt, der morsch war.* 'Recently the gardener showed the horsewoman a tree, that was decaying.' Structural representation of the genitive-reading, on the right(Chomsky, 1993, 1995, 2000). Example sentence of the genitive-reading: *Neulich hat der Gärtner der Reiterin den Baum gefällt, der morsch war.* 'Recently the gardener of the horsewoman chopped the tree, that was decaying.'

An approach proposed in Féry (2011) and adopted here is that every argument is contained in its own prosodic domain (φ-phrase or ι-phrase) by virtue of being a maximal projection (XP). Edges of these prosodic domains have been argued to be constructed by alignment with left or right edges of XPs

(Selkirk, 1986). The alignment of syntactic XPs with prosodic phrases is accounted for by generalized alignment constraints Align-Xp,r and Align-Xp,l introduced in (19), which demand the insertion of φ-boundaries to lexical XPs. For the dative case, NP1 and NP2 are organized in their own prosodic phrases with both, a left and a right φ-boundary between the two XPs. By contrast, in the genitive structure, NP2 is embedded in NP1 as a modifier to it. NP1 and NP2 thus represent one complex argument of V0. This complex argument forms its own prosodic phrase with an embedded φ-phrase wrapped around NP2.

Coming back to the dative structure. We argue that the first and second NP are contained in its own prosodic phrase. However, this phrasing violates a constraint that demands the entire XP to be included in one phonological phrase. This constraint is called Wrap-Xp (Truckenbrodt, 1999), introduced in (20). Note that Wrap-Xp is satisfied for the proposed phrasing of the genitive structure in providing one φ-phrase containing NP1 and its modifier NP2 within the same φ-phrase, phrased as (NP1 (NP2)φ)φ. Such phrasing would constitute a case of a recursive structure, where a prosodic domain of level n is part of a larger domain of the same type n. Recursivity in prosodic structure has been rejected in earlier works by Selkirk (1984) and Nespor and Vogel (1986). Ladd (1986), in turn, and also later accounts of Selkirk (1995) allowed for recursivity by providing arguments in favor of a violable constraint against recursive structure. This constraint prohibits recursive structures and is called Nonrecursivity (Selkirk, 1995; Truckenbrodt, 1999).

(39) Nonrecursivity
"Any two p-phrases that are not disjoint in extension are identical in extension."

The definition of Nonrecursivity in (39) claims that φ-phrases in recursive structures are similar in extension, meaning that a prosodic domain of one level may be included in a larger domain of the same level. The constraint thus penalizes recursive structures provided that the two elements of the recursive structure differ in extension. For the proposed recursive phrasing of the genitive structure Nonrecursivity would be violated, while it is not for the phrasal structure of the dative condition. Taken together, the interface constraints relevant for the present syntax-prosody mapping are Align-Xp,r, Align-Xp,l; Wrap-Xp and Nonrecursivity. These constraints will be applied to the abstract phonological structure in order to derive the correct prosodic phrase structure in an OT framework. This process is referred to as syntax-prosody mapping, operating at the surface level and outlined in the following section.

Evaluation of the dative phrase structure

	/ (NP1)ϕ	NP2 /	ALIGN-XP,R	WRAP-XP	NONREC	ALIGN-XP,L
☞	a. (NP1)ϕ	(NP2)ϕ		*		
	b. (NP1	NP2)ϕ	NP1!			NP2
	c. (NP1)ϕ	NP2)ϕ		*	NP2!	NP2

Figure 6.9: Tableau 40 [der GärtnerNP1 der ReiterinNP2]

The evaluation of the prosodic phrase structure of the dative condition is shown in Tableau (40). Tableau (40) shows two successive XPs as input, labelled as NP1 and NP2 in the input representation. Three possible output candidates with varying prosodic descriptions are presented, as generated by the function GEN. ALIGN-XP,R claims a φ-boundary after each lexical XP, and ALIGN-XP,L demands a φ-boundary before each XP. As for candidate a., the presence of φ-boundaries at the right and left edge of both NPs avoid violations of ALIGN-XP,R and ALIGN-XP,L. However, the phrasing in a. triggers violations of WRAP-XP since it provides no φ-phrase containing both, NP1 and NP1. Candidate b. does not meet the requirement of ALIGN-XP,R and is ruled out as a possible phrasing pattern. Candidate c. is a recursive structure that meets the demands of ALIGN-XP,R for NP1 and NP2 by showing a right edge of a φ-phrase at the right edge of NP1 and NP2. Similar to candidate a. it does violate WRAP-XP for the same reasons as does candidate a. However, NP2 violates NONRECURSIVITY and ALIGN-XP,L by showing no left φ-boundary between the two NPs. Therefore, NONRECURSIVITY can choose the correct phrasing in candidate a. over the incorrect recursive candidate c. In what follows, candidate a. represents the correct phrasing with both a left and a right φ-boundary between both NPs and, at the same time, indicates the dative phrasing as a nonrecursive structure, provided that NONRECURSIVITY is ranked below WRAP-XP which in turn is ranked below ALIGN-XP,R.

Evaluation of the genitive phrase structure

	/ (NP1	NP2 /	ALIGN-XP,R	WRAP-XP	NONREC	ALIGN-XP,L
	a. ((NP1)ɸ	NP2)ɸ			*	NP2!
	b. (NP1	NP2)ɸ	NP1!			NP2
	c. (NP1)ɸ	(NP2)ɸ		*!		
☞	d. (NP1	(NP2)ɸ)ɸ			*	

Figure 6.10: Tableau 41 [der GärtnerNP1 der ReiterinNP2]

Tableau 41 depicts the evaluation of the alternative prosodic phrase structure and shows that ALIGN-XP, R, ALIGN-XP, L together with WRAP-XP and NONRECURSIVITY derive a recursive structure as in candidate d., phrased as (NP1 (NP2)φ)φ. Again, ALIGN-XP,R demands a φ-boundary after each XP. This φ-boundary is not present in candidate b., and consequently, the phrasing is ruled out by ALIGN-XP,R. The φ-boundary called by ALIGN-XP,R is present in candidate a., candidate c. and candidate d. Consider first candidate a. It represents a recursive structure, where the outer φ-phrase fulfills the requirements of WRAP-XP, providing one φ-phrase containing both NPs. The inner φ-phrase obeys the demands of ALIGN-XP,R for the first NP by providing a right edge of a φ-phrase at the right edge of NP1. This inner right edge of φ however is not followed by a left edge of a φ-phrase. This structural formation is analyzed as a recursive prosodic structure. This recursive phrasing in a. thus violates NONRECURSIVITY and ALIGN-XP,L at the position of NP2 for reasons described above. Candidate c. presents a standard nonrecursive phrasing that one would otherwise expect for this structure. Recall, candidate c. does follow the demands of ALIGN-XP,R. Though, it is ruled out by WRAP-XP since it provides no joint φ-phrase wrapping NP1 and

NP2 together in one φ-phrase. Finally, candidate d. presents an alternative recursive structure which is likewise in accord with the demands of Align-Xp,r and Wrap-Xp. Further, it is favored over candidate a. by Align-Xp,l. Similar to candidate a., candidate d. violates Nonrecursivity. Notice, however, that in candidate d. the left internal boundary would correctly trigger NP2 as embedded within NP1. This constraint formation prefers the correct (( )φ)φ over the incorrect (( )φ)φ phrasal structure. As a consequence, candidate d. is the output that best harmonizes with the constraint hierarchy, as indicated by the pointing finger.

To summarize, the result of the syntax-prosody mapping is a prosodic phrase structure derived from a phonological structure via mapping syntactic information and prosodic information onto each other with recourse to a strict constraint ordering (Selkirk, 1986, 2005; Truckenbrodt, 1995, 1999). We saw that the derivation of a particular prosodic phrase structure can be accounted for by means of an OT grammar with four constraints that are ranked in the following order: Align-Xp, r » Wrap-Xp » Nonrecursivity » Align-Xp,l. These constraints are based on the assumptions and suggestions of previous studies claiming that they guide, among other constraints, the corresponding syntax-prosody mapping (Selkirk, 1986, 2005; Truckenbrodt, 1995, 1999). The application of these interface constraints on the abstract phonological structure derived a standard nonrecursive phrasing represented as (NP1)φ (NP2)φ and a recursive phrase structure, phrased as (NP1 (NP2)φ)φ.

Following the model in 6.2 above, the prosodic phrase structures are further analyzed and interpreted by listeners in a following processing step, referred to as *recognition*. This process assigns a meaning to the respective grammatical structures. As a result, the nonrecursive phrasing (NP1)φ/ι (NP2)φ would be linked with the dative condition, and the recursive phrase structure (NP1 (NP2)φ)φ with the genitive condition, in turn. The assignment of a particular meaning happens most likely via interpretative constraints operating at the recognition level.

6.5 Discussion

The proposed models apparently reflect the processing data obtained in the perception experiments involving φ-boundaries. At points of ambiguity, a φ-boundary is predicted for conditions with an NP1 duration exceeding 500ms (Tableau 25). This constitutes as clear case of a dative condition. In turn, no φ-boundary is predicted for inputs whose NP1 duration is 400ms or shorter (Tableau 26) which would then illustrate the alternative genitive condition. An input duration on NP1 lying inbetween 400ms and 500ms constitutes a case where listeners are no longer able to unambiguously map the phonetic signal onto either one of the two phonological structures, and consequently do not disambiguate the corresponding sentence fragments (Tableau 27). The present models' prediction of phonological phrase structure generation involving φ-boundaries rely on concrete duration values measured in the production experiment, that entered the corresponding perception experiments reported in the previous chapter 4. However, in order to make general predictions with respect to the phonetics-phonology mapping, one could alternatively operate with duration proportions among the relevant conditions, expressed in per cent for example, such that the maximal duration of a certain constituent, i.e. 100% (equals 500ms in the present case) guides the mapping onto a corresponding phonological structure, whereas a reduced duration of the same constituent, i.e. 80% (equals 400ms in the present case) leads to a mapping of the phonetic structure onto the alternative phonological structure in turn.

In case of an ι-boundary, deviating constraints as predicted for φ-boundaries become relevant: here, the correlate of the boundary tone preceding the choice point is predicted to be relevant. Specifically, the models predict an ι-boundary in case of a falling contour (Tableau 32), and no ι-boundary in case of a rising contour (Tableau 33). Therefore, the models are compatible with the general prevalence of falling accents preceding ι-boundaries that was observed in the perception experiment (cf. Experiment 1 in section 5.3, non-original conditions). The rejection of a rise on NP1 followed by an ι-boundary (Tableau 33, candidate d.) can be explained with recourse to the models' constraint *LH)ι (cf. (29), Tableau 33)) which ruled out the corresponding candidate. The competing candidate with a rise on NP1 in combination with the absence of an ι-boundary was established as the optimal interpretation (Tableau 33, candidate a.). That is, the presence or absence of a low ι-boundary is the result of the contour the listener is exposed to. Only a falling contour [HL] -> /(HL)ι/ -> ι-boundary (dative condition) leads to a low ι-boundary. Rising contours followed by a low ι-boundary are prohibited. Instead they are correlated either with or without a φ-boundary, depending on the duration of the critical constituent.

Generally, the proposed OT models account for the prosodic effects that were shown to influence syntactic ambiguity resolution in sentence comprehension. The assumption is that prosodic preferences at points of ambiguity reflect grammatical requirements, formulated as constraints operating at the phonetics-phonology interface. The proposed constraints that are thought to derive the phonetics-phonology mapping were independently motivated through perception experiments evaluating prosodic correlates on ambiguity resolution. In particular, the models assume phonetic correlates to be responsible for phrase structural generation that serve as input structure to the syntax-prosody mapping. Therefore, the present models are a grammatical device that assume prosodic constraints to act upon syntactic structure building. This is particularly obvious in chapter 4, Experiment 4 where participants opted more frequently to the genitive condition when local prosodic 'genitive cues' were inserted into a dative condition as compared to the unmanipulated dative condition presented in Experiment 1. These results stand is contrast to suggestions made by the mechanism underlying the parsing principle minimal attachment, assuming the minimal syntactic structure is adopted in the face of a syntactic ambiguity. The genitive reading, however, reflects a non-minimal attachment option. Yet, the finding that prosodic information influences parsing such that non-minimal interpretations would be pursued suggests that listeners do not automatically construct a purely syntactically driven minimal attachment structure but incorporate prosodic cues into initial parsing stages. Therefore, the present results question a default syntactic parsing mechanism in auditory language comprehension as formulated in serial parsing accounts (Frazier, 1979, 1987; Frazier & Clifton, 1996) or (Friederici, 1995, 2002). Such serial parsing accounts assume so-called two-stage mechanisms in order to comprehend a sentence. It is proposed that during the first-stage parse a simple syntactic structure is constituted only by consulting syntactic knowledge. In subsequent parsing stages other processing subsystems are called into play to establish a complete semantic, contextually integrated sentence representation. Does however semantic and/or contextual information conflicts with the initial syntactic parse, the syntactic structure is rejected and reconstructed after all.

Rather, the results of the present perception experiments and the resulting models provide evidence against serial parsing accounts and strongly suggest an interaction of syntax with other sources of information for initial structure building processes, as already pronounced by Zec & Inkelas (1990), among others. The view of such a bidirectional relation of syntax and other linguistic information has been proposed by a number of people, most prominently McClelland (1987); St. John & McClelland (1990) and led to the formulation of "interactive" constraint satisfaction models (Elman & McClelland, 1984; Trueswell et al., 1994; McRae et al., 1998b) assuming phonological, lexical, semantic, structural frequency information etc. to be used simultaneously during ambiguity resolution. Specifically, for the described results in the present work, an influence of phonetic/phonological information on syntactic structure building was validated, as φ- and ι-boundaries are interpreted by the parser as syntactic boundary markers, evidenced through the perception experiments and already claimed by Marcus & Hindle (1990). Therefore, prosodic boundary information can be seen as immediate prosodic influence on the computation of the respective syntactic structure and, therefore, the results are in accord with sentence processing models that ascribe prosody a role as input structure for the parser (Marcus & Hindle, 1990; Speer et al., 1996; Schafer, 1997).

We have shown that grammatical requirements transform the auditory input - the phonetic signal the listener is exposed to - into phonological objects, which are sequently interpreted as meaningful concepts to establish a possible mental representation. The preference of certain prosodic cues at points of ambiguity mirror the grammatical restrictions (constraints), respectively. Importantly, the relevant constraints correctly determing output forms, depend on the specific prosodic boundary size. Therefore, disambiguating the sentences presented in the present work involve the identification of the most relevant prosodic cue, most likely the most salient cue that signals a prosodic difference in the face of a syntactic ambiguity. Obviously, this cue identification takes place with maturity, growing language exposure. Perception studies with German 8-month-old infants on the question whether or not single prosodic boundary cues such as pitch change, prefinal lengthening and pause duration are sufficient for the identification of a φ-boundary in a sequence of three names with varying prosodic groupings ([Moni und Lilli und Manu] as opposed to [Moni und Lilli] [und Manu]) revealed that the combination of pitch change and prefinal lengthening is necessary for infants to complete the task successfully (Wellmann et al., 2012). In turn, our perception studies with adults on prosodic cue weighting indicate that just one relevant cue per boundary is sufficient. In this sense, our present results are in contrast with findings of Holzgrefe et al. (2012) who found that German adult listeners need a combination of two cues to reliably detect a prosodic phrase boundary in a prosodic judgement task. Holzgrefe et al. (2012) presented listeners sequences such as [a and b and c] and [a and b] [and c] and asked adults to interpret whether or not the auditory presented sequence contained an internal boundary after the second name. Sequences that contained either information of a pitch change or information of prefinal lengthening to indicate a boundary after the second name were interpreted as having no boundary. In contrast, sequences combining pitch change and prefinal lengthening within one utterance were reliably evaluated as containing an internal prosodic boundary. The diverging results with respect to adults' use of prosodic cues to identify a prosodic phrase boundary might indicate that the magnitude of the prosodic cue as such may influence listeners perception in identifying a prosodic boundary. It may likely that if the magnitude of each prosodic boundary cue is strong enough, a single prosodic cue is sufficient for listeners' perception of φ- and ι-phrase boundaries, as was shown in our experiments. With a diminished magnitude of each boundary marking cue, more than one cue might become necessary for listeners to identify the boundary. To summarize, if we evaluate the findings of German infants' and adults' use of prosodic cues, we can conclude that a change of prosodic boundary cues takes place from a holistic set required by infants younger than 6-8-month old over one for adults who require just a single boundary cue in dependence of the respective boundary type. The offered OT analyses in the present work captures this prosodic boundary related use of prosodic constraints for adults.

Comparing the results with non-intonational languages such as Akan or Hindi, different constraints that regulate the phonetics-phonology mapping might be involved. As for tone languages, numerous studies aim to scrutinize phrasal effects and their corresponding phonetic cues as well, see for instance Hyman (2003); Zerbian (2004). In Akan, representing a prototypical tone language where no lengthening effects preceding φ-boundaries are observable (Genzel, 2013), other prosodic parameters may serve to guide the phonetics-phonology mapping. It might be the case that the prevalent vowel harmony which is blocked preceding a φ-boundary (Kügler, 2012) may turn out to be a relevant boundary cue in perception. Additionally, glottal stops frequently occurring at the end of a word preceding a pause (Dolphyne, 1988) and pauses itself may serve as indicator of φ-boundaries in Akan as well. On the other hand, in Japanese, for example, a language exploiting a large number of intonational features (Beckman & Pierrehumbert, 1986a), the lowest level of phrasing which is referred to as Accentual Phrase (Beckman & Pierrehumbert, 1986a) is characterized by two tones, one being the high tone and the second one being the low boundary tone appearing at the end of every Accentual Phrase (except for cases where the L tone is present at the absolute beginning of the next Accentual Phrase (Beckman & Pierrehumbert, 1986a)). This suggests that tonal cues may serve to identify Accentual Phrases in perception. The Accentual Phrase itself is organized into a larger prosodic unit, which Beckmann & Pierrehumbert (1986) term intermediate phrase. Normally in Japanese, intermediate phrases can be delimited by a pause, by glottalisation, prefinal lengthening and the realization of the low boundary tone. Furthermore, the blocking of downstep at the position of the intermediate phrase boundary (Beckman & Pierrehumbert, 1986a) may be a prominent phonetic cue listeners might make recourse when encountering the intermediate phrase boundary. With regard to Hindi whose phrasal patterns were investigated by Patil et al. (2008), an analysis was proposed according to which each content word is phrased separately as one prosodic phrase (p-phrase). Each non-final p-phrase is associated with a low pitch accent and a high phrase boundary which is aligned with the right edge of the prosodic word (L*Hp). Final p-phrases have falling accents (H*Li). Sequences of p-phrases are organized in a downstep fashion which remains unaffected by focus, for instance. Focus, however, tends to be signaled by post-focal compression (Moore, 1965; Harnsberger & Judge, 1996; Patil et al., 2008). From the perspective of prosodic cues made available to listeners during perception, post focal compression may serve as prominent cue that reflects the focus structure of the sentence and thus serves as a cue to ambiguity resolution in Hindi (Kügler, 2011). Additionally, the observation that prefocal given constituents are not subject to compression may likewise be taken as a cue for listeners in determine the information structure of the sentence. Recall, that neither differences in pitch nor differences in prefinal lengthening are systematically employed by Hindi speakers to prosodically mark p-phrase boundaries of different embedded grouping conditions of three- and four names (Féry & Kentner, 2010). The present chapter indicated that certain phonetic features of the auditory input project a phonological phrase structure. In general, that mapping is guided by language specific constraints and in particular for German, the constraints relevant for the mapping are dependent on the size of the boundary.

6.6 Summary

In this chapter the results of a prosodic cue weighting on φ- and ι-phrase boundaries for the present case of the genitive-dative ambiguity were evaluated. Since both structural conditions appeared to be realized significantly different at points of ambiguity, and listeners' employed these prosodic effects for ambiguity resolution in perception, a processing account was proposed that modeled the perception of phonetic boundary correlates in an Optimality-Theoretic approach (Prince and Smolensky 1993/2004). Thereby the OT account makes direct recourse to grammatical constraints (cf. phonetic / phonological constraints) to determine parsing preferences at points of ambiguity. I have followed the assumption that the formation of an abstract phonological phrase structures result from a phonetics-phonology mapping, which in turn, serves as input structure for the corresponding syntax-prosody mapping yielding a prosodic phrase structure. The syntax-prosody mapping is guided via several interface constrains that are ordered according to a strict priority ranking to obtain the optimal phrase structure (Selkirk, 1986, 2005; Truckenbrodt, 1995, 1999). For the present case of a German case ambiguity with two successive NPs, the derivation of the dative and genitive phrase structure can be accounted for by means of four constraints which are ranked as follows: ALIGN-XP, R » WRAP-XP » NONRECURSIVITY » ALIGN-XP,L. Applying these constraints on the case ambiguity, the two NPs are phrased as (NP1)φ (NP2)φ in the dative condition, or alternatively, as a recursive phrase structure in the genitive condition, phrased as (NP1 (NP2)φ)φ. This analysis was shown to hold for structures involving either φ-phrase boundaries or ι-phrase boundaries at the position of NP1.

For the ambiguity at hand, the OT models allow us to derive testable processing predictions by assuming phonetic information to play a role during ambiguity resolution. It however, could in principle be extended to model linguistic phenomena in which syntactic, semantic, frequency, statistical effects, or discourse and referential contexts interact with phonetic and phonological ones. These models would be in accord with constraint based models like the one proposed by McRae et al. (1998b), among others, allowing all kinds of linguistic information to play a role in syntactic ambiguity resolution.

7

General conclusion and outlook

One of the central questions in psycholinguistic is understanding whether and how prosodic phrase boundaries are used to resolve syntactic ambiguities in sentence processing. The present work aimed to answer both, first, the effects of φ- and ι-boundaries on syntactic ambiguity resolution, and second, how the prosodic correlates of the auditory input are taken for the phonetic-phonology mapping in order to attain a meaningful sentence interpretation.

With regard to the first aim, we investigated locally syntactic ambiguities involving either φ- or ι-phrase boundaries in German and the structural preference that listeners have, based on the prosodic content. The experiments described in this work show that German listeners exploit both types of prosodic phrase boundaries to resolve local syntactic ambiguities, that however, their disambiguation altered by the presence or absence of prosodic cues correlated with the corresponding boundary. Specifically, the perception data revealed that the phonetically measured prosodic correlates of each prosodic boundary such as pitch accents, boundary tones, deaccentuation and durational properties do not contribute to ambiguity resolution in equal measure. Rather, it is the case that listeners rely primarily on *prefinal lengthening* as a correlate of phrasing in the vicinity of φ-phrase boundaries, while at the level of the ι-phrase boundary, boundary tones serve as phrasal cues. This way the results of the present work take account of the as yet missing information on individual contributions of prosodic correlates on listeners' disambiguation of syntactically ambiguous sentences in German. It further implies that the question of how German listeners resolve syntactic ambiguities cannot simply be attributed to the presence or absence of prosodic correlates. The interpretation of the phrasal structure rather depends on a more general picture of cohesion between prosodic correlates and prosodic boundary sizes.

With respect to the second aim, the processing models proposed in the present work describe a specific phonetic-phonology mapping in the vicinity of both phrase boundaries. It is assumed that auditory sentence processing proceeds in several successively organized steps, during which listeners transform overt phonetic forms into language specific abstract surface forms. This process is referred to as phonetic-phonology mapping in the present work. Perceptual evidence resulting from the experiments of the present work suggest that the phonetic-phonology mapping is guided by the above mentioned boundary related prosodic correlates. The resulting abstract phonological structure is subjected to the syntax-prosody mapping, in turn. The outcome of the presented perception experiments are modulated in an Optimality-Theoretic framework. The offered OT-models are grounded on the assumption that single prosodic correlates are used by listeners as a signal to syntax in sentence processing. This is in line with studies arguing that the prosodic phrase structure determines the syntactic parse (Cutler et al., 1997; Warren et al., 1995; Pynte & Prieur, 1996; Snedeker & Trueswell, 2003; Kjelgaard & Speer, 1999), to name just a few.

Certainly, since we concentrated exclusively on the role of prosodic correlates on sentence disambiguation by examining one type of syntactic ambiguity, several questions remain open.

First, since the observations reported in the present work are based on one particular syntactic ambiguity, further research is needed to determine if the tendencies hold across a larger set of syntactic ambiguities, such as prepositional phrase attachment, adverb attachment, or direct object versus sentence complement noun phrase attachment. In that context, the issue of prosodic boundary strength could be focus of examination. Our data suggest that the prosodic cue-use is boundary size dependent. Whether such a strategy constitutes the scaffold on which German listeners construct a cue-weighting or whether their cue-use would shift as a result of different sentence structures could be subjected to further research. Such studies would complement or strengthen the established cue weighting for German.

Second, the lexical influence remained unexplored. In order to exclude other factors than prosodic parameters that might affect sentence disambiguation, a prosodic cue weighting using delexicalised stimuli would be expedient, for example, either by low-pass filtering the stimuli to make them unintelligible while preserving the intonational contour, or by using pseudo words to assure an exclusive impact of prosodic information detached from lexical information.

Another possible research endeavor refers to participants' individual sensitivity to prosodic cues. It it most likely that there are gradual differences between listeners with respect to the strength they rely on either structural or prosodic information.

Also, the interplay of phrasing and accentuation is untouched in the present work. Accent structure plays a role in sentence processing as well. A possible hypothesis is that if a constituent is prominent by accentuation, it will be perceived to be phrased separately. Non-prominent constituents are integrated into adjacent prosodic phrases. That is, a break between a non-prominent constituent and a neighboring constituent will be less informative with respect to syntactic structure than the same break between two prominent constituents. It is promising to test this hypothesis with the ambiguity at hand and examine the interplay of phrasing and accentuation on sentence disambiguation. When the ambiguous second NP is deaccented, listeners should integrate this constituent into the previous phrase and consequently tend to interpret it as a genitive NP. When both, NP1 and NP2 are accented, this should lead more often to the dative interpretation. This can be tested within a cross-model sentence completion task where listeners will be presented controlled auditory sentence fragments up to and including the ambiguous sequence and will be asked to chose the sentence continuation they think is appropriate. Apart from accentuation, the strength of the prosodic boundary between NP1 and NP2 will be controlled for.

The experiments reported here examine individual prosodic influences on auditory sentence disambiguation, in particular the interaction between prosodic cues and boundary strength during sentence processing and fill a gap by providing controlled experimental studies that examined the individual contribution of each prosodic correlate on listeners disambiguation of syntactically ambiguous sentences in German. Clearly, more work is needed to fully reveal the underlying processing mechanisms of auditory sentence disambiguation, in particular by acknowledging information from other linguistic domains as well. However, the current results provide important groundwork for future research by identifying which specific prosodic renderings significantly influence auditory sentence disambiguation.

Bibliography

Aasland, W. & Baum, S. (2003). Temporal parameters as cues to phrasal boundaries: a comparison of processing by left-hemisphere-damaged and right-hemisphere-damaged individuals. *Brain and Language*, *87*, 385–399.

Augurzky, P. & Schlesewsky, M. (2010). Prosodic phrasing and transitivity in head-final sentence comprehension: Erp evidence from german ambiguous dps. *Processing and Producing Head-Final Structures*.

Avesani, C. & Vayra, M. (2005). Accenting, deaccenting and information strcture in italien dialogue. In *SIGdial6-2005*, (pp. 19–24).

Bates, D. & Sarkar, D. (2013). lme4: Linear mixed-effects models using s4 classes. r package version 0.9975-11.

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure: Evidence for cue trading relations. *Journal of Memory and Language*, *30*, 644–663.

Beckman, M. & Edwards, J. (1990). Lengthening and shortening and the nature of prosodic constituency. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 152–178). Kingston, J. and Beckman, M.

Beckman, M. & Pierrehumbert, J. (1986a). Intonational structure in japanese and english. *Phonology Yearbook*, *3*, 255–309.

Beckman, M. & Pierrehumbert, J. (1986b). Japanese prosodic phrasing and intonation synthesis. *Proceedings of the 24th ACL, New York*, 173–180.

Benzmüller, R. & Grice, M. (1998). The nuclear accentual fall in the intonation of standard german. *ZAS Papers in Linguistics 11*, 79–98.

Berkovits, R. (1994). Durational effects in final lengthening, gapping, and contrastive stress. *Language and Speech*, *37*(3), 237–250.

Birch, S. & Clifton, C. (1995). Focus, accent, and argument structure: effects on language comprehension. *Language and Speech*, *38*, 365–391.

Boersma, P. (1998). *Functional Phonology*. PhD thesis, University of Amsterdam. The Hague: Holland Academic Graphics.

Boersma, P. (2000). The ocp in the perception grammar. *Rutgers Optimality Archive 435*.

Boersma, P. (2009). *Cue constraints and their interactions in phonological perception and production*, (pp. 55–110). Berlin: Mouton De Gruyter.

Boersma, P., Escudero, P., & Hayes, R. (2003). Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. In *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, 3-9 August 2003*, (pp. 1013–1016).

Boersma, P. & Hamann, S. (2009). *Introduction: models of phonology in perception*, (pp. 1–24). Berlin: Mouton De Gruyter.

Boersma, P. & Weenink, D. (2013). Praat-doing phonetics by computer. *online resource: http://www.praat.org*.

Bolinger, D. (1972). Accent is predictable (if you're not a mind reader). *Language*, *48*, 633–644.

Brants, S., Dipper, S., Eisenberg, P., Hansen-Schirra, S., König, E., Lezius, W., Rohrer, C., Smith, G., & Uszkoreit, H. (2004). Tiger: Linguistic interpretation of a german corpus. *Research on Language and Computation*, *2*, 597–619.

Brants, S., Dipper, S., Hansen, S., Lezius, W., & Smith, G. (2002). The tiger treebank. In *Proceedings of the workshop on treebanks and linguistic theories year*, (pp. 24–41). Citeseer.

Bruce, G. (1977). *Swedish word accent in sentence perspective*, volume 12. Gleerup, Lund.

Büring, D. (1997). *The 49th Bridge Accent*. Berlin: Mouton De Gruyter.

Butcher, A. (1981). Aspects of the speech pause: Phonetic correlates and communicative functions. kiel: Institut für phonetik, universität kiel. *Kiel: Institut für Phonetik, Universität Kiel*.

Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America*, *120*(3), 1589–1599.

Chen, A. (2007). Language-specificity in the perception of continuation intonation. In C. Gussenhoven & T. Riad (Eds.), *Tones and tunes II: Phonetic and behavioural studies in word and sentence prosody* (pp. 107–142). Berlin: Mouton De Gruyter.

Chen, M. (1987). The syntax of xiamen tone sandhi. *Phonology Yearbook*, *4*, 109–149.

Choi, H. (2003). Prosody-induced acoustic variation in english stop consonants. In *Proceedings of the 15th International Congress of Phonetic Science*, (pp. 2661–2664).

Chomsky, N. (1993). *A minimalist program for linguistic theory*, volume 20, (pp. 1–52). Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: the framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by Step: essays on Minimalist syntax in honour of Howard Lasnik* (pp. 89–115). Cambridge, MA: MIT Press.

Clark, J. & Yallop, C. (1995). *An Introduction to Phonetics and Phonology*.

Clements, G. (1978). Tone and syntax in ewe. In D. J. N. Napoli (Ed.), *Elements of tone, stress, and intonation* (pp. 21–99). Washington, D.C.: Georgetown University Press.

Clifton, C., Speer, S., & Abney, S. (1991). Parsing arguments: Phrase structure and arguments structure as determinants of initial parsing decisions. *Journal of Memory and Language*, *30*, 251–271.

Collier, R., de Pijper, J., & Sanderman, A. (1993). Perceived prosodic boundaries and their phonetic correlates. In Morgan Kaufman Publishers, Plainsboro, N. J. U. (Ed.), *Proceeding of the ARPA Workshop on Human Language Technology*, (pp. 341–345).

Collins, B. & Mees, I. M. (1981). The phonetics of english and dutch. In *Brill, Leiden*.

Cooper, W. & Paccia-Cooper, J. (1980). Syntax and speech. *Cambridge, MA: Harvard University Press*.

Cruttenden, A. (1986). *Intonation*. PhD thesis, Cambridge: CUP.

Cutler, A., Dahan, D., & Donselaar, W. v. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*, 141–201.

de Pijper, J. & Sanderman, A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustical Society of America*, *96*(4), 2037–2047.

De Vincenzi, M. & Job, R. (1993). Some observations on the universality of the late closure strategy. *Journal of Psycholinguistic Research*, *22*(2), 189–206.

De Vincenzi, M. & Job, R. (1995). An investigation of late closure: The role of syntax, thematic structure, and pragmatics in initial and final interpretation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*(5), 1303–1321.

Dehé, N. (2014). *Parentheticals in Spoken English: The Syntax-Prosody Relation*. Cambridge University Press.

Delattre, P. (1966). A comparison of syllable length conditioning among languages. *IRAL-International Review of Applied Linguistics in Language Teaching*, *4*(1-4), 183–198.

Denes, P. (1955). Effects of duration on the perception of voicing. *Journal of the Acoustical Society of America*, *27*, 761–764.

Di Cristo, A. & Hirst, D. J. (1997). *L' accentuation non-emphatique en francais: stratégies et paramètres*. Paris: L'Harmattan, Paris.

D'Imperio, M. & Gili-Favela, B. (2004). How many levels of phrasing? evidence from two varieties of italien. In Local, J., Ogden, R., & Temple, R. (Eds.), *Phonetic Interpretation, Papers in Laboratory Phonology VI*, (pp. 130–144). Cambridge: Cambridge University Press.

Dolphyne, F. A. (1988). The akan (twi-fante) language: Its sound systems and tonal structure. *Ghana: University Press.*

Elfner, E. (2012). *Syntax-prosody interactions in Irish*. PhD thesis, University of Massachusetts, Amherst.

Elman, J. L. & McClelland, J. L. (1984). Speech perception as a cognitive process: The interactive activation model. In N. Lass (Ed.), *Speech and Language*, volume 10. New York: Academic Press.

Escudero, P. (2005). *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization*. PhD thesis, Utrecht University.

Escudero, P. & Boersma, P. (2003). Modelling the perceptual development of phonological contrasts with optimality theory and the gradual learning algorithm. In Arunachalam, S., Kaiser, E., & Williams, A. (Eds.), *Proceedings of the 25th Annual Penn Linguistics Colloquium. Penn Working Papers in Linguistics*, volume 8, (pp. 71–85).

Fant, G. & Kruckenberg, A. (1996). On the quantal nature of speech timing. In *Spoken Language*, (pp. 2044–2047).

Fant, G. & Kruckenberg, A. (1998). Prominence and accentuation. acoustical correlates. In *Proceedings of Fonetik -98, The Swedish Phonetic Conference, Stockholm University* (pp. 142–145). Branderud, P. and Traunmüller, H.

Fernald, A. & McRoberts, G. (1996). Prosodic bootstrapping: a critical analysis of the argument and the evidence. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 365–388). Mahwah, NJ: Lawrence Earlbaum Associates.

Ferreira, F. (1988). *Planning and timing in sentence production: The syntax-to-phonology conversion*. PhD thesis, Unpublished dissertation, University of Massachusetts, Amherst, MA.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, *30*, 210–233.

Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, *2*, 233–253.

Ferreira, F. & Henderson, J. (1990). The use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 555–569.

Féry, C. (1993). German intonational patterns. *Linguistische Arbeiten, Niemeyer*.

Féry, C. (2010a). Indian languages as intonational "phrase languages". In I. Hasnain & S. Chaudhury (Eds.), *Problematizing language Studies. Festschrift for Rama Agnihotri* (pp. 288–312). Delhi: Aakar Books.

Féry, C. (2010b). *Prosodische Phrasen und Intonation: der Bezug zwischen Syntax und Phonologie am Beispiel des Deutschen*. Linguisten-Band 9 der Japanischen Gesellschaft für Germanisten.

Féry, C. (2010c). Recursion in prosodic structure. *Phonological Studies 2010*, *13*.

Féry, C. (2011). German sentence accents and embedded prosodic phrases. *Lingua*, *121*, 1906–1922.

Féry, C. (2012). Die annotation der deutschen intonation: Wie kann gtobi die syntax und die semantik integrieren? *Linguistische Berichte*, *2012*(229), 39–97 (59).

Féry, C. & Ishihara, S. (2010). How focus and givenness shape prosody. In M. Zimmermann & C. Féry (Eds.), *Information Structure. Theoretical, Typological, and Experimental Perspectives* (pp. 36–63). Oxford: Oxford University Press.

Féry, C., Kaiser, E., Hörnig, R., Weskott, T., & Kliegl, R. (2009). *Perception of intonational contours on given and new referents: a completion study and an eye-movement experiment*, (pp. 235–266). Berlin: Mouton De Gruyter.

Féry, C. & Kentner, G. (2010). The prosody of grouping names in german and hindi. In *Proceedings of Speech Prosody 2010, Chicago*.

Féry, C. & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, *36*, 680–703.

Féry, C. & Samek-Lodovici, V. (2006). Focus projection and prosodic prominence in nested foci. *Language*, *82*(1), 131–150.

Féry, C. & Schubö, F. (2011). Hierarchical prosodic structures in the intonation of center-embedded relative clauses. *The Linguistic Review*, *27*, 289–313.

Féry, C. & Stoel, R. (2006). Gradient perception of intonation. In G. Fanselow, C. Féry, M. Schlesewsky, & R. Vogel (Eds.), *Gradience in Grammar: Generative Perspectives* (pp. 145–166). Oxford: Oxford University Press.

Flemming, E. (1995). *Auditory Representations in Phonology*. PhD thesis, Stanford University.

Fletcher, J. (1991). Rhythm and final lengthening in french. *Journal of Phonetics*, *19*, 193–212.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In *The Handbook of Phonetic Sciences, Wiley-Blackwell, 523-602.* (pp. 523–602). Wiley-Blackwell.

Frazier, L. (1979). On comprehending sentences: Syntactic parsing strategies. *ETD Collection for University of Connecticut*.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 559–585). Hillsdale, NJ: Erlbaum.

Frazier, L. & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.

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

Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambigous sentences. *Cognitive Neuropsychology*, *14*(178-210).

Friederici, A. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259–281.

Friederici, A. (2002). Towards a neural basis of auditory language processing. *Trends in Cognitive Science*, *6*, 78–84.

Friedrich, C. (2003). *Prosody and spoken word recognition: Behavioral and ERP correlates*. PhD thesis, MPI Series in Cognitive Neuroscience, 36, Leipzig.

Fry, D. (1955). Duration and intensity as physical correlates of stress. *Journal of the Acoustic Society of America*, *27*, 318–336.

Gee, J. P. & Grosjean, F. (1983). Performance structure: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*, 411–458.

Genzel, S. (2013). *Lexical and post-lexical tones in Akan*. PhD thesis, Universität Potsdam.

Gibson, E., Pearlmutter, N., Canseco-Gonzales, E., & Hickock, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, *59*, 23–59.

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech*, 103–113.

Grabe, E. (1998). *Comparative Intonational Phonology. English and German*. PhD thesis, Nijmegen University.

Grice, M. & Baumann, S. (2002). Deutsche intonation und gtobi. *Linguistische Berichte*, *191*, 267–298.

Grice, M., Baumann, S., & Benzmüller, R. (2005). German intonation in autosegmental-metrical phonology. In S.-A. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 55–83). Oxford: Oxford University Press.

Grice, M., Baumann, S., & Jagdfeld, N. (2009). Tonal association and derived nuclear accents - the case of downstepping contours in german. *Lingua*, *119*(6), 881–905.

Grice, M., Reyelt, M., Benzmüller, R., Mayer, J., & Batliner, A. (1996). Consistency in transcription and labelling of german intonation with gtobi. *Proc. ICSLP Philadelphia*.

Grosjean, F. (1983). How long is the sentence? prediction and prosody in the on-line processing of language. *Linguistics*, *21*(3), 501–530.

Grosjean, F. & Collins, M. (1979). Breathing, pausing and reading. *Phonetica*, *36*, 98–114.

Grosjean, F. & Hirt, C. (1996). Using prosody to predict the end of sentences in english and french: Normal and brain-damaged subjects. *Language and Cognitive Processes*, *11*(1), 107–134.

Gussenhoven, C. (1983). Testing the reality of focus domains. *Language and Speech*, *26*, 61–80.

Gussenhoven, C. (1984). *On the Grammar and Semantics of Sentence Accents*. Dordrecht, Holland; Cinnaminson, N.J., USA: Foris Publication.

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation. Research surveys in linguistics.* Cambridge: Cambridge University Press.

Halle, M. & Isardi, W. (1995). General properties of stress and metrical structure. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 403–443). Oxford: Basil Blackwell.

Halle, M. & Vergnaud, J. (1987). An essay on stress.

Harnsberger, J. & Judge, J. (1996). Pitch range and focus in hindi. *The Journal of the Acoustical Society of America*, *99*(4), 2493.

Hayes, B. (1995). Metrical stress theory: Principles and case studies. *University of Chicago Press*.

Heldner, M. (2001). *Focal accent - f0 movements and beyond*. PhD thesis, Umea University.

Henderson, A. I. & Nelms, S. (1980). Relative salience of intonation fall and pause as cues to the perceptual segmentation of speech in an unfamiliar language. *Journal of Psycholinguistic Research*, *9*, 147–159.

Hirst, D. & Di Cristo, A. (1998). *Intonation Systems: A survey of Twenty Languages*. Cambridge University Press.

Hirst, D. J. (1999). The symbolic coding of duration and timing: An extension to the intsint system. *Proceedings Eurospeech 99*, 1639–1642.

Hoequist, C., J. (1983). Durational correlates of linguistic rhythm categories. *Phonetica*, *40*, 19–31.

Hogan, J. & Rozsypal, A. (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America*, *67*, 1764–1771.

Holzgrefe, J., Schröder, C., Höhle, B., & Wartenburger, I. (2012). Processing of prosodic boundary cues as revealed by event-related brain potentials. *Poster Presented at the 13th Conference on Laboratory Phonology (LabPhon 13), Stuttgart*.

Horne, M., Strangert, E., & Heldner, M. (1995). Prosodic boundary strength in swedish: Final lengthening and silent interval duration. In *Proceedings of the 13th International Congress of Phonetic Science*, (pp. 170–173).

House, A. & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*, *25*, 105–113.

House, D. (1990). *Tonal perception in speech*. PhD thesis, Lund, Travaux de l'Institute de Linguistique de Lund 24.

House, D., Hermes, D., & Beaugendre, F. (1998). Perception of tonal rises and falls for accentuation and phrasing in swedish. In *Proceedings of the International Congress of Phonetic Science* (pp. 2799–2802). Mannell. R. and Robert-Ribes, J.

Hruska, C., Alter, K., Steinhauer, K., & Steube, A. (2001). Misleading dialogues: human's brain reaction to prosodic information. In C. Cavé, I. Guaïtella, & S. Santi (Eds.), *Oralité et gesturalité: Interactions et comportements multimodaux dans la communication* (pp. 425–430). Paris: L'Harmattan.

Hyman, L. M. (2003). African languages and phonological theory. *GLOT International*, *7*(6), 153–163.

Ishihara, S. (2007). Major phrase, focus intonation and multiple spell-out (map, fi, mso). *The Linguistic Review*, *24*, 137–167.

Ito, J. & Mester, A. (2009). The extended prosodic word. In J. Grijzenhout & B. Kabak (Eds.), *Phonological Domains: Universals and Deviations* (pp. 135–194). Mouton de Gruyter: Interface Explorations Series.

Ito, J. & Mester, A. (2011). Recursive prosodic phrasing in japanese. In T. Borowsky, S. Kawahara, T. Shinya, & M. Sugahara (Eds.), *Prosody matters: Essays in honor of Elisabeth Selkirk*. London: Equinox Publishers.

Jacobs, J. (1993). Integration. In M. Reis (Ed.), *Wortstellung und Informationsstruktur* (pp. 63–116). Tübingen: Niemeyer.

Johnson, E. K. & Seidl, A. (2008). Clause segmentation by 6-month-olds: A crosslinguistic perspective. *Infancy*, *13*, 440–455.

Jun, S. A. (2004). Prosody in sentence processing: Korean vs. english. *UCLA Working Papers in Phonetics*, *104*, 26–45.

Jun, S.-A. (2005). *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford University Press.

Kager, R. (1995). The metrical theory of word stress. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 367–402). Oxford: Basil Blackwell.

Kahnemuyipour, A. (2004). *The syntax of sentential stress*. PhD thesis, University of Toronto.

Kisseberth, C. & Abasheikh, M. (1974). Vowel length in chimwi:ni: A case study of the role of grammar in phonology. In *Papers from the Parasession on Natural Phonology, Chicago Linguistic Society*, volume 10 (pp. 193–209). Chicago Linguistic Society, University of Chicago, Chicago, Ill.

Kjelgaard, M. & Speer, S. (1994). Prosodic structure influences syntactic attachments during the course of sentence comprehension. Presented to the Annual CUNY Sentence Processing Conference, New York.

Kjelgaard, M. & Speer, S. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic ambiguity. *Journal of Memory and Language*, *40*, 153–194.

Klatt, D. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics*, *3*, 129–140.

Klatt, D. (1976). Linguistic uses of segmental duration in english: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Kohler, K. J. (1983). Prosodic boundary signals in german. *Phonetica*, *40*, 89–134.

Kratzer, A. & Selkirk, E. (2007). Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review*, *24*, 93–135.

Krivokapic, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, *35*, 162–179.

Kügler, F. (2007). *The Intonational Phonology of Swabian and Upper Saxon*. PhD thesis, Tübingen: Niemeyer (Linguistische Arbeiten 515).

Kügler, F. (2011). Post-focal pitch register compression as a prosodic cue for disambiguation in hindi. *Poster presented at Speaking of prosody - A symposium in conjunction with the 60th birthday of Professor Merle Horne - Lund*.

Kügler, F. (2012). Segmental and prosodic cues for phrasing in akan. *Talk held at the Workshop of the Syntax-Phonology Interface from a Cross-linguistic perspective, ZAS, Berlin*.

Kuzla, C., Cho, T., & Ernestus, M. (2007). Prosodic strengthening of german fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, *35*, 301–320.

Ladd, B. (1984). Declination: a review and some hypotheses. *Phonology Yearbook*, *1*, 53–74.

Ladd, B. (1986). Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook*, *3*, 311–340.

Ladd, B. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Ladd, B. & Campbell, D. (1991). Theories of prosodic structure: evidence from syllable duration. In *Proceedings of the 12th International Congress of Phonetic Science*, (pp. 290–293).

Ladd, D. R. (1988). Declination "reset" and the hierarchical organization of utterances. *Journal of the Acoustical Society of America*, *84*, 530–544.

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *The Journal of the Acoustical Society of America*, *53*, 380.

Lehiste, I., Olive, J., & Streeter, L. (1976). Role of duration in disambiguating syntactically ambiguous sentences. *The Journal of the Acoustical Society of America*, *60*, 1199.

Li, W. & Yang, Y. (2009). Perception of prosodic hierarchical boundaries in mandarin chinese sentences. *Neuroscience*, *158*(4), 1416–1425.

Liberman, A. & Douglas, H. W. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, *4*, 187–196.

Liberman, M. (1975). *The Intonational System of English*. PhD thesis, MIT.

Liberman, M. & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R. T. Oehrle (Eds.), *Language, sound, structure: studies in phonology presented to Morris Halle by his teacher and students* (pp. 157–233). Cambridge: MIT Press.

Liberman, M. & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, *8*, 249–336.

Macmillan, N. A. & Creelman, C. D. (2005). *Detection theory: a user's guide*. Cambridge University Press.

Marcus, M. & Hindle, D. (1990). Description theory and intonation boundaries. In G. Altman (Ed.), *Cognitive Models of Speech Processing: Psycholinguistics and Computational Perspectives* (pp. 483–512). Cambridge, MA: MIT Press.

Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., & Lee, C. S. (1992). Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology*, *45*, 73–87.

Martin, E. (1970a). Toward and analysis of subjective phrase structure. *Psychological Bulletin*, *74*, 153–166.

Martin, J. (1970b). On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior*, *9*(1), 75–78.

Mayer, J. (1997). *Intonation und Bedeutung. Aspekte der Prosodie-Semantik-Schnittstelle im Deutschen.* PhD thesis, Universität Stuttgart.

McCarthy, J. & Prince, A. (1993). *Prosodic Morphology I: Constraint Interaction and Satisfaction, ms.* University of Massachusetts, Amherst, and Rutgers University.

McClelland, J. L. (1987). The case of interactionism in language processing. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 3–36). Hillsdale, NJ: Erlbaum.

McQueen, J. & Cutler, A. (1997). Cognitive processes in speech perception. In W. J. Hardcastle & J. Laver (Eds.), *The Handbook of Phonetic Science* (pp. 566–586). Oxford: Blackwell.

McRae, K., Spivey-Knowlton, M., & Tanenhaus, M. (1998a). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998b). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312.

Mietz, A., Toepel, U., Ischebeck, A., & Alter, K. (2008). Inadequate and infrequent are not alike: Erps to deviant prosodic patterns in spoken sentence comprehension. *Brain and Language*, *104*(2), 159–169.

Millotte, S., René, A., Wales, R., & Christophe, A. (2008). Phonological phrase boundaries constrain the online syntactic analysis of spoken sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 874–885.

Mo, Y. (2008). Duration and intensity as perceptual cues for naive listeners' prominence and boundary perception. *Proceedings of the 4th Conference on Speech Prosody, Campinas, Brazil*, 739–742.

Möbius, B. (1993). Ein quantitatives modell der deutschen intonation: Analyse und synthese der grundfrequenzverläufe. *Tübingen: Niemeyer.*

Moore, R. (1965). *A study of Hindi intonation*. PhD thesis, University of Michigan.

Neppert, J. & Pétursson, M. (1986). Elemente einer akustischen Phonetik. *Hamburg: H. Buske Verlag*.

Nespor, M. & Vogel, I. (1986). Prosodic Phonology. *Mouton de Gruyter*.

Nittrouer, S., Manning, C., & Mayer, G. (1993). The perceptual weighting of acoustic cues change with linguistic experience. *Journal of the Acoustical Society of America*, *94*, 1865–1865.

Nittrouer, S. & Miller, M. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America*, *101*, 2253–2266.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in english. *Journal of the Acoustical Society of America*, *54*, 1235–1247.

O'Malley, M., Kloker, D., & Dara-Abrams, B. (1973). Recovering parentheses from spoken algebraic expressions. *IEEE Transactions on Audio and Electroacoustics*, *21*(2), 217–220.

Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., & Vasishth, S. (2008). Focus, word order, and intonation in hindi. *Journal of South Asian Linguistics*, *1*, 53–70.

Peters, J. (2006). *Intonation deutscher Regionalsprachen*. Walter de Gruyter: Berlin New York.

Peterson, G. & Lehiste, I. (1960). Duration of syllable nuclei in english. *Journal of the Acoustical Society of America*, *32*, 693–703.

Pierrehumbert, J. & Beckman, M. (1988). Japanese tone structure. *Cambridge: MIT Press*.

Pierrehumbert, J. & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in Communication* (pp. 271–311). Cambridge, MA: MIT Press.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. PhD thesis, MIT.

Pierrehumbert, J. B. (1981). Synthesizing intonation. *Journal of the Acoustical Society of America*, *70*(4), 985–995.

Poser, W. J. (1984). *The phonetics and phonology of tone and intonation in Japanese*. PhD thesis, MIT, Cambridge, Massachussetts.

Price, P., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in disambiguation. *Journal of the Acoustical Society of America*, *90*, 2956–2970.

Prieto, P., Chilin, S., & Holly, N. (1996). Pitch downtrends in spanish. *Journal of Phonetics*, *24*, 445–473.

Prieto, P., D'Imperio, M., & Fivela Gili, B. (2005). Pitch accent alignment in romance: primary and secondary associations with metrical structure. *Language and Speech*, *4*, 359–396.

Prince, A. (1983). Relating to the grid. *Linguistic Inquiry*, *14*, 19–100.

Prince, A. & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar. RuCCS Technical Report 2*. Rutgers University, Piscateway, NJ: Rutgers University Center for Cognitive Science.

Pynte, J. & Prieur, B. (1996). Prosodic breaks and attachment decisions in sentence parsing. *Language and Cognitive Processes*, *11*, 165–191.

Rayner, K. & Pollatsek, A. (1989). *The psychology of reading*. Routledge.

Reyelt, M., Grice, M., Benzmüller, R., Mayer, J., & Batliner, A. (1996). Prosodische etikettierung des deutschen mit tobi. *Proc. KONVENS 96*, 290–299.

Sanderman, A. & Collier, R. (1997). Prosodic phrasing and comprehension. *Language and Speech*, *40*, 391–409.

Schafer, A. (1997). *Prosodic Parsing: The role of prosody in sentence comprehension*. PhD thesis, University of Massachusetts, Amherst.

Scheer, T. (2012). How phonological is intonation? *Paper presented at the Jahrestagung der Deutschen Gesellschaft für Sprache (DGFS), Frankfurt am Main 7-9 March*.

Schmitz, M. (2009). The perception of clauses in 6- and 8-month-old german-learning infants: Influence of pause duration and the natural pause hierarchy. *Universitätsverlag Potsdam*.

Schneider, K. (2012). *The German boundary tones: Categorical Perception, Perceptual Magnets, and the Perceptual Reference Space*. PhD thesis, University of Stuttgart.

Schneider, K. & Lintfert, B. (2003). Categorical perception of boundary tones in german. *Proceedings of the 15th International Congress of Phonetic Sciences (Barcelona)*, 631–634.

Schneider, K. & Möbius, B. (2005). Perceptual magnet effect in german boundary tones. *Proceedings of Interspeech/Eurospeech, Lisbon*, 41–44.

Schütze, C. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. University of Chicago Press.

Schütze, C. & Gibson, E. (1999). Argumenthood and english prepositional phrase attachment. *Journal of Memory and Language*, *40*, 409–431.

Scobbie, J. M., Gibbon, F., Hardcastle, W. J., & Fletcher, P. (1996). Covert contrast as a stage in the acquisition of phonetics and phonology. *QMC Working papers in Speech and Language Science*, *1*, 43–62.

Scott, D. (1982). Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America*, *71*, 996–1007.

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, *57*, 24–48.

Seidl, A. & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, *11*, 596–606.

Selkirk, E. (1984). Phonology and syntax: the relation between sound and structure. *MIT, Cambridge, MA*.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, *3*, 371–405.

Selkirk, E. (1990). On the nature of prosodic constituency. *Papers in Laboratory Phonology*, *1*, 179–200.

Selkirk, E. (1995). Sentence prosody: Intonation, stress and phrasing. In J. Goldsmith (Ed.), *Handbook of Phonological Theory* (pp. 550–569). Blackwell, Oxford.

Selkirk, E. (1996). The prosodic structure of function words. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: bootstrapping from speech to grammar in early acquisition* (pp. 187–214). Mahwah, NJ: Erlbaum.

Selkirk, E. (2000). The interaction of constraints on prosodic phrasing. In M. Horne (Ed.), *Prosody: Theory and Experiment* (pp. 231–262). Dordrecht: Kluwer.

Selkirk, E. (2005). Comments on intonational phrasing. In V. M. Frota, S. & M. Freitas (Eds.), *Prosodies* (pp. 11–58). Selkirk, Elisabeth O. (2005). Comments on intonational phrasing. In Sonia Frota, Marina Mouton de Gruyter.

Selkirk, E. (2009). On clause and intonational phrase in japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu, 136*.

Selkirk, E. (2011). The syntax-phonology interface. In J. Goldsmith, J. Riggle, & A. Yu (Eds.), *The Handbook of Phonological Theory, 2nd edition*. Oxford: Blackwell.

Shattuck-Hufnagel, S. & Turk, A. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*, 193–247.

Snedeker, J. & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*(1), 103–130.

Sosa, J. (1999). *La entonación del espanol: Su estructura fónica, variabilidad y dialectología*. Madrid: Cátedra.

Speer, S., Kjelgaard, M., & Dobroth, K. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, *25*(2), 249–271.

St. John, M. F. & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, *46*, 217–257.

Steinhauer, K. (2011). Combining behavioral measures and brain potentials to study categorical prosodic boundary perception and relative boundary strength. In *Proceedings of the 17th International Congress of Phonetic Science*, (pp. 1898–1901).

Steinhauer, K., Alter, K., & Friederici, A. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*, 191–196.

Strangert, E. (1990). Where do pauses occur in text read alound? In *Proc. Twelfth Scand. Conf. of Linguistics, Reykjavik*.

Strangert, E. (1991). Pausing in texts read aloud. In *Proceedings of the 12th International Congress of Phonetic Science*, (pp. 238–241).

Streeter, L. (1978). Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America*, *64*, 1582–1592.

t′ Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation*. Cambridge University Press.

Tabain, M. (2003). Effects of prosodic boundary on /ac/ sequences: Articulatory results. *Journal of the Acoustical Society of America*, *113*, 2834–2849.

Terken, J. & Collier, R. (1991). Syntactic influences on prosody. In Y. Tokhura, Vatikiotis-Bateson, & Y. Sagisaki (Eds.), *Speech perception, production and linguistic structure* (pp. 427–438). Oxford: IOS Press.

Truckenbrodt, H. (1995). *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. PhD thesis, Massachusetts Institute of Technology, Cambridge, Mass.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *30*(2), 219–255.

Truckenbrodt, H. (2002). Upstep and embedded register levels. *Phonology*, *19*, 77–120.

Truckenbrodt, H. (2004). Final lowering in non-final position. *Journal of Phonetics*, *32*, 313–348.

Truckenbrodt, H. (2005). A short report on intonation phrase boundaries in german. *Linguistische Berichte*, *203*, 273–296.

Trueswell, J. C., Tanenhaus, M., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285–318.

Turk, A. & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in american english words. *Journal of Phonetics*, *27*, 171–206.

Turk, A. & White, L. (1999). Structural influences on accentual lengthening in english. *Journal of Phonetics*, *27*(2), 171–206.

Uhmann, S. (1991). Fokusphonologie: eine analyse deutscher intonationskonturen im rahmen der nicht-linearen phonologie. *Tübingen: Niemeyer.*

Vaissière, J. (1983). Language-independent prosodic features. In A. Cutler & D. Ladd (Eds.), *Prosody: Models and measurements* (pp. 53–66). Springer.

Van Petten, C. & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time-course of meaning activation. *Journal of Memory and Language*, *26*, 188–208.

Vayra, M. & Fowler, C. (1992). Declination of supralaryngeal gestures in spoken italian. *Phonetica*, *49*(1), 48–60.

Venditti, J., Jun, S. A., & Beckman, M. (1996). Prosodic cues to syntactic and other linguistic structure in japanese, korean and english. In *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 287–311). Mahwah, NJ: Lawrence Earlbaum Associates.

Volskaya, N. (2003). Virtual and real pauses at clause and sentence boundaries. *Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona*.

von Essen, O. (1956). Grundzüge der hochdeutschen satzintonation. *Ratingen: Henn.*

Wagner, M. (2005). *Prosody and recursion*. PhD thesis, Massachusetts Institute of Technology, Cmbridge, Mass.

Wagner, M. & Watson, D. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, *25*(7-9), 905–945.

Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes*, *10*, 457–486.

Wellmann, C., Holzgrefe, J., Truckenbrodt, H., Wartenburger, I., & Höhle, B. (2012). How each prosodic boundary cue matters: Evidence from german infants. *Frontiers in Psychology*, *3*, doi: 10.3389/fpsyg.2012.00580.

Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91*, 1707–1717.

Willems, N. (1982). English intonation from a dutch point of view. In *Foris Publications*.

Wunderlich, D. (1988). Der ton macht die melodie - zur phonologie der intonation des deutschen. In H. Altmann (Ed.), *Intonationsforschungen* (pp. 1–40). Tübingen: Niemeyer.

Xinting, Z. (2011). Cue-weighting in the perception of prosodic boundaries in chinese and english. *Mid-Phon 17, Mid-Continental Phonetics and Phonology Conference, University of Illinois at Urbana-Champaign*.

Yang, L. (2007). Duration, pauses, and the temporal structure of mandarin spontaneous speech. In Trouvain, J. & Barry, W. J. (Eds.), *Proceedings of the 16th International Congress of Phonetic Science*, (pp. 1289–1292).

Zec, D. & Inkelas, S. (1990). Prosodically constrained syntax. In S. Inkelas & D. Zec (Eds.), *The phonology-syntax connection* (pp. 365–378). University of Chicago Press.

Zellner, B. (1994). Pauses and the temporal structure of speech. In E. Keller (Ed.), *Fundamentals of Speech Synthesis and Speech Recognition*. Chichester: John Wiley.

Zerbian, S. (2004). Phonological phrases in xhosa (southern bantu). *ZAS Papers in Linguistics*, *37*, 71–99.

8
Appendix A

Sentence material used in production Experiment 1 (Chapter 4)

1. Kürzlich hat der Opa der Tänzerin ein Haus {entdeckt, vererbt}, das sehr zerfallen war.
2. Zögernd hat der Vater der Schwimmerin das Camp {besucht, erlaubt}, in dem viel trainiert wird.
3. Mittwoch hat der Vater der Bänkern ein Fax {verschickt, geschickt}, das geheim bleiben sollte.
4. Schliesslich hat der Nachbar der Mieterin den Brief {gekriegt, gebracht}, der lange erwartet wurde.
5. Montag hat der Lehrer der Künstlerin das Buch {gelesen, gegeben}, das neu erschienen war.
6. Samstag hat der Pfarrer der Sängerin den Chor {geleitet, vorgestellt}, der sich neu gegründet hat.
7. Unlängst hat der Maler der Zauberin das Bild {entwendet, gefunden}, das er schon immer haben wollte.
8. Gerade hat der Bauer der Wärterin das Kalb {entdeckt, vererbt}, das schon lange verschwunden war.
9. Ungern hat der Opa der Schülerin den Schnaps {geleert, spendiert}, der sehr stark war.
10. Mehrmals hat der Richter der Klägerin die Mail {geschickt, verschickt}, die den Prozessverlauf schildert.
11. Neulich hat der Gärtner der Reiterin den Baum {gezeigt, gefällt}, der morsch war.
12. Mittags hat der Pfleger der Rentnerin den Saft {umgekippt, aufgedrängt}, der nicht schmeckte.
13. Endlich hat der Bruder der Forscherin das Geld {ausgegeben, überwiesen}, von dem er gesprochen hat.
14. Abends hat der Kellner der Maurerin ein Bier {geleert, spendiert}, das schön kalt war.

15. Gestern hat der Trainer der Siegerin den Preis {entwendet, gefunden}, den sie gewonnen hat.
16. Diesmal hat der Förster der Pächtern ein Rind {geschenkt, gesehen}, das sehr jung war.
17. Erstmals hat der Onkel der Kaiserin das Gut {gezeigt, besucht}, das ihm vererbt wurde.
18. Schliesslich hat der Winzer der Bäuerin das Fass {aufgedrängt, umgekippt}, in dem teurer Wein war.
19. Morgens hat der Vetter der Fleischerin das Lamm {gesehen, geschenkt}, das verkauft werden sollte.
20. Freitag hat der Schwager der Sängerin den Text {gebracht, gekriegt}, den er schon gesucht hat.

Sentence material and contexts used in production Experiment 2 (Chapter 5)

1. Der Anwalt der Klägerin war bereits eingetroffen. Im vorderen Abschnitt des Raumes herrschte heftiges Treiben denn dort saß der Anwalt der Klägerin auf einer Bank.
2. In der letzten Reihe des Gerichtssaales war noch ein Platz frei. Dort saß der Anwalt. Der Klägerin blieb keine andere Möglichkeit als in den vorderen Reihen Platz zu nehmen.
3. Unsere Oberin hatte Familienbesuch. Ihre Geschwister waren gekommen. Gestern fuhr der Bruder der Oberin wieder nach Wien.
4. Die Klosteranlage soll einen Anbau bekommen. Alle Ordensbrüder wurden gebeten mitzuhelfen. Nur einer half für 2 Tage. Gestern fuhr der Bruder. Der Oberin missfiel die Nachricht.
5. Die Justizvollzugsanstalt ist in heller Aufruhr. Eine Wärterin wurde ermordet. Gestern floh der Mörder der Wärterin über die Grenze.
6. Das Gefängnis ist in heller Aufruhr. Gedeckt von einer Wärterin konnte ein Mörder tagelang im Frauentrakt ein- und ausgehen. Gestern floh der Mörder. Der Wärterin wurde sofort gekündigt.
7. Gestern sprang die Schwimmerin vom 3-Meter Turm. Ihr Vater schaute zu. Heute sprang der Vater der Schwimmerin ins Wasser.
8. In der letzten Woche war das Wasser im Schwimmbecken noch wärmer. Die Schwimmerin sprang jeden Tag vom Turm ins Schwimmbecken und ihr Vater schaute ihr dabei zu. Heute sprang der Vater. Der Schwimmerin war das Wasser zu kalt.
9. Die Schwägerin hatte von Anfang an Probleme mit ihren Dritten Zähnen. Heute half der Zahnarzt der Schwägerin ihre Dritten Zähne einzusetzten.
10. Bisher wurde die Schwägerin durch ihren Mann beim Einsetzten der Dritten Zähne unterstüzt. Heute half der Zahnarzt. Der Schwägerin fiel ein Stein vom Herzen.

11. Die Fahrerin des VOLVOs konnte das Auto nicht mehr starten und es musste angeschoben werden. Gott sei Dank war sie nicht allein im Auto. Nun schob der Partner der Fahrerin das Auto wieder an.
12. Die Fahrerin konnte das Auto nicht mehr starten. Zunächst probierte sie alleine den Wagen wieder anzuschieben, aber sie hatte zu wenig Kraft. Nun schob der Partner. Der Fahrerin war das sehr recht.
13. In der Kanzlei herrschte für alle Alkoholverbot, doch ausgerechnet beim Fahrdienst fiel das nie auf. Immer trank der Fahrer der Anwältin sein Glas Wein.
14. Immer trank der Fahrer. Der Anwältin musste jedes Mal ein Taxi gerufen werden. Als es ihr zu viel wurde, hat sie ihn gefeuert.
15. Jahrelang hing Claras Überleben als Malerin von den Zuwendungen des alten Grafen ab. Kürzlich starb der Gönner der Künstlerin überraschend. Clara war am Ende.
16. Claras Karriere als Malerin hing vom Geld ihres Mannes und den Zuwendungen eines Gönners, des alten Grafen ab. Kürzlich starb der Gönner. Der Künstlerin entging die Hälfte ihrer Einnahmen.
17. Der neue Postgaul überraschte alle. Er schlug aus und lief einige Meter über den Hof. Zögernd rief der Bote der Herzogin das Pferd zurück.
18. Der Herzog sagte zu den Jünglingen, dass der erste, der rufe, seine Tochter bekäme. Zögernd rief der Bote. Der Herzogin missfiel die Nachricht.
19. Die Tierpflegerin stand mit einem Käfig in dem grossen Vogelhaus und lockte ihren Vogel um ihn einzufangen. Endlich flog der Vogel der Pflegerin in den Käfig.
20. Die Heimangestellten kümmerte sich wochenlang um den kranken Vogel. Aber dieser wollte nicht fliegen. Endlich flog der Vogel. Der Pflegerin fiel ein Stein vom Herzen.
21. Es ist Winter geworden, doch Adis Schützling übt weiter die Aufschläge. Neulich fror der Trainer der Sportlerin auf der Bank.
22. Neulich fror der Trainer. Der Sportlerin hingegen war sehr warm.
23. Im Pisa-Lernstudio herrscht heute Abend Flaute. Nur ein Mädchen ist gekommen und der alte Herr Kurth ist müde. Ständig träumt der Lehrer der Schülerin während der Nachhilfe.
24. Was ist heute nur mit Herrn Kurth los? Ständig träumt der Lehrer. Der Schülerin platzt bald der Kragen.
25. Das Problem im Obstverkauf sind nicht die Preise, sondern die Kunden. Nehmen wir zum Beispiel Frau Nandkes kleinen Laden und ihren Stammgast, Herrn Sommer. Immer isst der Kunde der Händlerin im Obstladen, ohne zu bezahlen.
26. Die Feinkosthändlerin schaut am Ende der Woche das Filmmaterial ihrer neuen Videoüberwachungsanlage an und freut sich, darauf ihren attraktiven Stammkunden zu erkennen. Doch was ist das? Immer isst der Kunde. Der Händlerin vergeht das Lachen.

27. In seinem zweiten Lehrjahr beschäftigt sich Hans nun mit anspruchsvolleren Aufgaben. Häufig plant der Lehrling der Gärtnerin die Bepflanzung der Gartenanlagen.

28. Seit er im zweiten Lehrjahr ist, wagt sich Hans auch selbst an die Gartengestaltung. Häufig plant der Lehrling. Der Gärtnerin gefällt sein Engagement.

29. In unserer neuen Wohnung ist es eigentlich sehr still. Nur nachts weint der Junge der Nachbarin manchmal im Schlaf.

30. Der Film «Die Nachbarin« ist wirklich ergreifend. Immer wieder denke ich an die Szene mit dem kranken Kind. – Nachts weint der Junge. Der Nachbarin gehen die Ideen aus, um ihren Sohn zu beruhigen.

31. Kaum jemand glaubt Maria, dass ihr Ex-Mann sie mit der Trompete fast erschlagen hätte, doch ihr Nachbar hat die Tat beobachtet. Heute redet der Zeuge der Sängerin im Gerichtssaal.

32. Die Anklage präsentiert einen fast blinden Mann, der vorgibt, die Tat beobachtet zu haben. Heute redet der Zeuge. Der Sängerin ist das recht; sie erhofft sich dadurch den Frei-spruch.

9

Appendix B

TIGER-Corpus request for the experimental conditions

Genitive condition request:
NPnom . NPgen embedded . NPacc followed by the verb
output: 13

```
tiger:pos=/(V[AM]FINIKOUS)/ & #7 . #1 & tiger:cat="NP" & tiger:morph=/Nom.*/ & #1
_l_ #2 & tiger:cat="NP" & #1 > #3 & #1 _r_ #3 & tiger:morph=/Gen.*/ & #3 _l_ #4 &
tiger:cat="NP" & #3 . #5 & tiger:morph=/Acc.*/ & # 6 _l_ #5 & tiger:pos=/V.*/ & #5 .
#8
```

Dative condition request:
NPnom . NPdat . NPacc followed by the verb
output: 14

```
tiger:pos=/(V[AM]FINIKOUS)/ & #7 . #1 & tiger:cat="NP" & tiger:morph=/Nom.*/ & #1
_l_ #2 & tiger:cat="NP" & #1 . #3 & tiger:morph=/Dat.*/ & #3 _l_ #4 & tiger:cat="NP"
& #3 . #5 & tiger:morph=/Acc.*/ & #6 _l_ #5 & tiger:pos=/V.*/ & #5 . #8
```

10
Appendix C

Durational manipulation of the sentence material tested in perception Experiment 3 (Chapter 4)

manipulated words	con	original duration	compressed/ expanded to
|der Gärtner| |der Reiterin| |den Baum|	DAT	650 700 470	555 725 460
	GEN	460 750 450	555 725 460
|der Pfleger| |der Rentnerin| |den Saft|	DAT	780 550 420	605 635 455
	GEN	430 720 490	605 635 455
|der Bruder| |der Forscherin| |das Geld|	DAT	580 640 350	495 695 405
	GEN	410 750 460	495 695 405
| der Kellner| |der Maurerin| |ein Bier|	DAT	640 540 340	535 570 415
	GEN	430 600 490	535 570 415
|der Trainer| |der Siegerin| |den Preis|	DAT	690 440 460	565 560 505
	GEN	440 680 550	565 560 505
|der Pförster| |der Pächterin| |ein Rind|	DAT	660 480 400	540 545 425
	GEN	420 610 450	540 545 425
|der Onkel| |der Kaiserin| |das Gut|	DAT	640 600 380	485 630 410
	GEN	330 660 440	485 630 410
|der Winzer| |der Bäuerin| |das Fass|	DAT	600 550 370	485 560 405
	GEN	370 570 440	485 560 405
|der Vetter| |der Fleischerin| |das Lamm|	DAT	680 630 430	498 610 440
	GEN	317 590 450	498 610 44
|der Schwager| |der Sängerin| |den Text|	DAT	730 490 420	585 525 435
	GEN	440 560 450	585 525 435
|der Opa| |der Tänzerin| |ein Haus|	DAT	530 570 400	465 650 425
	GEN	400 730 450	465 650 425
|der Vater| |der Schwimmerin| |das Camp|	DAT	650 530 390	540 620 400
	GEN	430 710 410	540 620 400
|der Vater| |der Bänkerin| |ein Fax|	DAT	620 510 530	540 570 520
	GEN	460 630 510	540 570 520
|der Nachbar| |der Mieterin| |den Brief|	DAT	630 490 370	520 535 410
	GEN	410 580 450	520 535 410

<table>
<tr><th>manipulated words</th><th>con</th><th colspan="3">original duration</th><th colspan="3">compressed/ expanded to</th></tr>
<tr><td>|der Lehrer| |der Künstlerin| |das Buch|</td><td>DAT</td><td>580</td><td>600</td><td>400</td><td>490</td><td>655</td><td>435</td></tr>
<tr><td></td><td>GEN</td><td>400</td><td>710</td><td>470</td><td>490</td><td>655</td><td>435</td></tr>
<tr><td>|der Pfarrer| |der Sängerin| |den Chor|</td><td>DAT</td><td>570</td><td>510</td><td>400</td><td>490</td><td>580</td><td>485</td></tr>
<tr><td></td><td>GEN</td><td>410</td><td>650</td><td>570</td><td>490</td><td>580</td><td>485</td></tr>
<tr><td>|der Maler| |der Zauberin| |das Bild|</td><td>DAT</td><td>640</td><td>590</td><td>460</td><td>525</td><td>610</td><td>465</td></tr>
<tr><td></td><td>GEN</td><td>410</td><td>630</td><td>470</td><td>525</td><td>610</td><td>465</td></tr>
<tr><td>|der Bauer| |der Wärterin| |das Kalb|</td><td>DAT</td><td>490</td><td>520</td><td>410</td><td>410</td><td>645</td><td>405</td></tr>
<tr><td></td><td>GEN</td><td>330</td><td>770</td><td>400</td><td>410</td><td>645</td><td>405</td></tr>
<tr><td>|der Opa| |der Schülerin| |den Schnaps|</td><td>DAT</td><td>590</td><td>540</td><td>370</td><td>450</td><td>540</td><td>430</td></tr>
<tr><td></td><td>GEN</td><td>310</td><td>540</td><td>490</td><td>450</td><td>540</td><td>430</td></tr>
<tr><td>|der Richter| |der Klägerin| |die Mail |</td><td>DAT</td><td>590</td><td>600</td><td>390</td><td>470</td><td>660</td><td>405</td></tr>
<tr><td></td><td>GEN</td><td>350</td><td>720</td><td>420</td><td>470</td><td>660</td><td>405</td></tr>
</table>

Pause duration manipulation in ms of the sentence material tested in perception Experiment 3 (Chapter 4)

sent pair	manipulated pause	condition	original duration	changed to
1.	|P1|	DAT	0.13	0.075
	|P1|	GEN	0.02	0.075
2.	|P1|	DAT	0.04	0.035
	|P1|	GEN	0.03	0.035
3.	|P1|	DAT	0.10	0.07
	|P1|	GEN	0.04	0.07
4.	|P1|	DAT	0.06	0.04
	|P1|	GEN	0.02	0.04
5.	|P1|	DAT	0.08	0.055
	|P1|	GEN	0.03	0.055
6.	|P1|	DAT	0.07	0.05
	|P1|	GEN	0.03	0.05
7.	|P1|	DAT	0.09	0.055
	|P1|	GEN	0.02	0.055
8.	|P1|	DAT	0.06	0.04
	|P1|	GEN	0.02	0.04
9.	|P1|	DAT	0.06	0.04
	|P1|	GEN	0.02	0.04
10.	|P1|	DAT	0.10	0.06
	|P1|	GEN	0.02	0.06
11.	|P1|	DAT	0.10	0.10
	|P1|	GEN	0.10	0.10
12.	|P1|	DAT	0.10	0.065
	|P1|	GEN	0.03	0.065
13.	|P1|	DAT	0.12	0.075
	|P1|	GEN	0.03	0.075
14.	|P1|	DAT	0.07	0.035
	|P1|	GEN	0.00	0.035
15.	|P1|	DAT	0.08	0.055
	|P1|	GEN	0.03	0.055
16.	|P1|	DAT	0.12	0.075
	|P1|	GEN	0.03	0.075
17.	|P1|	DAT	0.09	0.06
	|P1|	GEN	0.03	0.06
18.	|P1|	DAT	0.09	0.06
	|P1|	GEN	0.03	0.06
19.	|P1|	DAT	0.40	0.20
	|P1|	GEN	0.01	0.20
20.	|P1|	DAT	0.12	0.07
	|P1|	GEN	0.02	0.07

Word duration manipulation in ms of the dative sequence tested in perception Experiment 4 (Chapter 4)

<table>
<tr><th>manipulated words</th><th>con</th><th colspan="3">original duration</th><th colspan="3">compressed/ expanded to</th></tr>
<tr><td>|der Gärtner| |der Reiterin| |den Baum|</td><td>DAT</td><td>650</td><td>700</td><td>470</td><td>460</td><td>750</td><td>450</td></tr>
<tr><td></td><td>GEN</td><td>460</td><td>750</td><td>450</td><td>460</td><td>750</td><td>450</td></tr>
<tr><td>|der Pfleger| |der Rentnerin| |den Saft|</td><td>DAT</td><td>780</td><td>550</td><td>420</td><td>460</td><td>750</td><td>450</td></tr>
<tr><td></td><td>GEN</td><td>460</td><td>750</td><td>450</td><td>460</td><td>750</td><td>450</td></tr>
<tr><td>|der Bruder| |der Forscherin| |das Geld|</td><td>DAT</td><td>580</td><td>640</td><td>350</td><td>410</td><td>750</td><td>460</td></tr>
<tr><td></td><td>GEN</td><td>410</td><td>750</td><td>460</td><td>410</td><td>750</td><td>460</td></tr>
<tr><td>| der Kellner| |der Maurerin| |ein Bier|</td><td>DAT</td><td>640</td><td>540</td><td>340</td><td>430</td><td>600</td><td>490</td></tr>
<tr><td></td><td>GEN</td><td>430</td><td>600</td><td>490</td><td>430</td><td>600</td><td>490</td></tr>
<tr><td>|der Trainer| |der Siegerin| |den Preis|</td><td>DAT</td><td>690</td><td>440</td><td>460</td><td>440</td><td>680</td><td>550</td></tr>
<tr><td></td><td>GEN</td><td>440</td><td>680</td><td>550</td><td>440</td><td>680</td><td>550</td></tr>
<tr><td>|der Pförster| |der Pächterin| |ein Rind|</td><td>DAT</td><td>660</td><td>480</td><td>400</td><td>420</td><td>610</td><td>450</td></tr>
<tr><td></td><td>GEN</td><td>420</td><td>610</td><td>450</td><td>420</td><td>610</td><td>450</td></tr>
<tr><td>|der Onkel| |der Kaiserin| |das Gut|</td><td>DAT</td><td>640</td><td>600</td><td>380</td><td>330</td><td>660</td><td>440</td></tr>
<tr><td></td><td>GEN</td><td>330</td><td>660</td><td>440</td><td>330</td><td>660</td><td>440</td></tr>
<tr><td>|der Winzer| |der Bäuerin| |das Fass|</td><td>DAT</td><td>600</td><td>550</td><td>370</td><td>370</td><td>570</td><td>440</td></tr>
<tr><td></td><td>GEN</td><td>370</td><td>570</td><td>440</td><td>370</td><td>570</td><td>440</td></tr>
<tr><td>|der Vetter| |der Fleischerin| |das Lamm|</td><td>DAT</td><td>680</td><td>630</td><td>430</td><td>317</td><td>590</td><td>450</td></tr>
<tr><td></td><td>GEN</td><td>317</td><td>590</td><td>450</td><td>317</td><td>590</td><td>450</td></tr>
<tr><td>|der Schwager| |der Sängerin| |den Text|</td><td>DAT</td><td>730</td><td>490</td><td>420</td><td>440</td><td>560</td><td>450</td></tr>
<tr><td></td><td>GEN</td><td>440</td><td>560</td><td>450</td><td>440</td><td>560</td><td>450</td></tr>
<tr><td>|der Opa| |der Tänzerin| |ein Haus|</td><td>DAT</td><td>530</td><td>570</td><td>400</td><td>400</td><td>730</td><td>450</td></tr>
<tr><td></td><td>GEN</td><td>400</td><td>730</td><td>450</td><td>400</td><td>730</td><td>450</td></tr>
<tr><td>|der Vater| |der Schwimmerin| |das Camp|</td><td>DAT</td><td>650</td><td>530</td><td>390</td><td>430</td><td>710</td><td>410</td></tr>
<tr><td></td><td>GEN</td><td>430</td><td>710</td><td>410</td><td>430</td><td>710</td><td>410</td></tr>
<tr><td>|der Vater| |der Bänkerin| |ein Fax|</td><td>DAT</td><td>620</td><td>510</td><td>530</td><td>460</td><td>630</td><td>510</td></tr>
<tr><td></td><td>GEN</td><td>460</td><td>630</td><td>510</td><td>460</td><td>630</td><td>510</td></tr>
<tr><td>|der Nachbar| |der Mieterin| |den Brief|</td><td>DAT</td><td>630</td><td>490</td><td>370</td><td>410</td><td>580</td><td>450</td></tr>
<tr><td></td><td>GEN</td><td>410</td><td>580</td><td>450</td><td>410</td><td>580</td><td>450</td></tr>
<tr><td>|der Lehrer| |der Künstlerin| |das Buch|</td><td>DAT</td><td>580</td><td>600</td><td>400</td><td>400</td><td>710</td><td>470</td></tr>
<tr><td></td><td>GEN</td><td>400</td><td>710</td><td>470</td><td>400</td><td>710</td><td>470</td></tr>
<tr><td>|der Pfarrer| |der Sängerin| |den Chor|</td><td>DAT</td><td>570</td><td>510</td><td>400</td><td>410</td><td>650</td><td>570</td></tr>
<tr><td></td><td>GEN</td><td>410</td><td>650</td><td>570</td><td>410</td><td>650</td><td>570</td></tr>
<tr><td>|der Maler| |der Zauberin| |das Bild|</td><td>DAT</td><td>640</td><td>590</td><td>460</td><td>410</td><td>630</td><td>470</td></tr>
<tr><td></td><td>GEN</td><td>410</td><td>630</td><td>470</td><td>410</td><td>630</td><td>470</td></tr>
<tr><td>|der Bauer| |der Wärterin| |das Kalb|</td><td>DAT</td><td>490</td><td>520</td><td>410</td><td>330</td><td>770</td><td>400</td></tr>
<tr><td></td><td>GEN</td><td>330</td><td>770</td><td>400</td><td>330</td><td>770</td><td>400</td></tr>
<tr><td>|der Opa| |der Sülerin| |den Schnaps|</td><td>DAT</td><td>590</td><td>540</td><td>370</td><td>310</td><td>540</td><td>490</td></tr>
<tr><td></td><td>GEN</td><td>310</td><td>540</td><td>490</td><td>310</td><td>540</td><td>490</td></tr>
<tr><td>|der Richter| |der Klägerin| |die Mail |</td><td>DAT</td><td>590</td><td>600</td><td>390</td><td>350</td><td>720</td><td>420</td></tr>
<tr><td></td><td>GEN</td><td>350</td><td>720</td><td>420</td><td>350</td><td>720</td><td>420</td></tr>
</table>

Pause duration manipulation in ms of the dative sequence tested in perception Experiment 4 (Chapter 4)

sent pair	manipulated pause	condition	original duration	changed to
1.	|P1|	DAT	0.13	0.02
	|P1|	GEN	0.02	0.02
2.	|P1|	DAT	0.04	0.03
	|P1|	GEN	0.03	0.03
3.	|P1|	DAT	0.10	0.04
	|P1|	GEN	0.04	0.04
4.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
5.	|P1|	DAT	0.08	0.03
	|P1|	GEN	0.03	0.03
6.	|P1|	DAT	0.07	0.03
	|P1|	GEN	0.03	0.03
7.	|P1|	DAT	0.09	0.02
	|P1|	GEN	0.02	0.02
8.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
9.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
10.	|P1|	DAT	0.10	0.02
	|P1|	GEN	0.02	0.02
11.	|P1|	DAT	0.10	0.10
	|P1|	GEN	0.10	0.10
12.	|P1|	DAT	0.10	0.03
	|P1|	GEN	0.03	0.03
13.	|P1|	DAT	0.12	0.03
	|P1|	GEN	0.03	0.03
14.	|P1|	DAT	0.07	0.00
	|P1|	GEN	0.00	0.00
15.	|P1|	DAT	0.08	0.03
	|P1|	GEN	0.03	0.03
16.	|P1|	DAT	0.12	0.03
	|P1|	GEN	0.03	0.03
17.	|P1|	DAT	0.09	0.03
	|P1|	GEN	0.03	0.03
18.	|P1|	DAT	0.09	0.03
	|P1|	GEN	0.03	0.03
19.	|P1|	DAT	0.40	0.01
	|P1|	GEN	0.01	0.01
20.	|P1|	DAT	0.12	0.02
	|P1|	GEN	0.02	0.02

Pause duration manipulation in ms of the dative sequence tested in perception Experiment 5 (Chapter 4)

sent pair	manipulated pause	condition	original duration	changed to
1.	|P1|	DAT	0.13	0.02
	|P1|	GEN	0.02	0.02
2.	|P1|	DAT	0.04	0.03
	|P1|	GEN	0.03	0.03
3.	|P1|	DAT	0.10	0.04
	|P1|	GEN	0.04	0.04
4.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
5.	|P1|	DAT	0.08	0.03
	|P1|	GEN	0.03	0.03
6.	|P1|	DAT	0.07	0.03
	|P1|	GEN	0.03	0.03
7.	|P1|	DAT	0.09	0.02
	|P1|	GEN	0.02	0.02
8.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
9.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
10.	|P1|	DAT	0.10	0.02
	|P1|	GEN	0.02	0.02
11.	|P1|	DAT	0.10	0.10
	|P1|	GEN	0.10	0.10
12.	|P1|	DAT	0.10	0.03
	|P1|	GEN	0.03	0.03
13.	|P1|	DAT	0.12	0.03
	|P1|	GEN	0.03	0.03
14.	|P1|	DAT	0.07	0.00
	|P1|	GEN	0.00	0.00
15.	|P1|	DAT	0.08	0.03
	|P1|	GEN	0.03	0.03
16.	|P1|	DAT	0.12	0.03
	|P1|	GEN	0.03	0.03
17.	|P1|	DAT	0.09	0.03
	|P1|	GEN	0.03	0.03
18.	|P1|	DAT	0.09	0.03
	|P1|	GEN	0.03	0.03
19.	|P1|	DAT	0.40	0.01
	|P1|	GEN	0.01	0.01
20.	|P1|	DAT	0.12	0.02
	|P1|	GEN	0.02	0.02

Segmental duration manipulation of the sentence material tested in perception Experiment 6 (Chapter 4)

manipulated words	con	original duration			compressed/ expanded to		
|der Gärtner| |der Reiterin| |den Baum|	DAT	650	700	470	555	725	460
	GEN	460	750	450	555	725	460
|der Pfleger| |der Rentnerin| |den Saft|	DAT	780	550	420	605	635	455
	GEN	430	720	490	605	635	455
|der Bruder| |der Forscherin| |das Geld|	DAT	580	640	350	495	695	405
	GEN	410	750	460	495	695	405
| der Kellner| |der Maurerin| |ein Bier|	DAT	640	540	340	535	570	415
	GEN	430	600	490	535	570	415
|der Trainer| |der Siegerin| |den Preis|	DAT	690	440	460	565	560	505
	GEN	440	680	550	565	560	505
|der Pförster| |der Pächterin| |ein Rind|	DAT	660	480	400	540	545	425
	GEN	420	610	450	540	545	425
|der Onkel| |der Kaiserin| |das Gut|	DAT	640	600	380	485	630	410
	GEN	330	660	440	485	630	410
|der Winzer| |der Bäuerin| |das Fass|	DAT	600	550	370	485	560	405
	GEN	370	570	440	485	560	405
|der Vetter| |der Fleischerin| |das Lamm|	DAT	680	630	430	498	610	440
	GEN	317	590	450	498	610	440
|der Schwager| |der Sängerin| |den Text|	DAT	730	490	420	585	525	435
	GEN	440	560	450	585	525	435
|der Opa| |der Tänzerin| |ein Haus|	DAT	530	570	400	465	650	425
	GEN	400	730	450	465	650	425
|der Vater| |der Schwimmerin| |das Camp|	DAT	650	530	390	540	620	400
	GEN	430	710	410	540	620	400
|der Vater| |der Bänkerin| |ein Fax|	DAT	620	510	530	540	570	520
	GEN	460	630	510	540	570	520
|der Nachbar| |der Mieterin| |den Brief|	DAT	630	490	370	520	535	410
	GEN	410	580	450	520	535	410
|der Lehrer| |der Künstlerin| |das Buch|	DAT	580	600	400	490	655	435
	GEN	400	710	470	490	655	435
|der Pfarrer| |der Sängerin| |den Chor|	DAT	570	510	400	490	580	485
	GEN	410	650	570	490	580	485
|der Maler| |der Zauberin| |das Bild|	DAT	640	590	460	525	610	465
	GEN	410	630	470	525	610	465
|der Bauer| |der Wärterin| |das Kalb|	DAT	490	520	410	410	645	405
	GEN	330	770	400	410	645	405
|der Opa| |der Schülerin| |den Schnaps|	DAT	590	540	370	450	540	430
	GEN	310	540	490	450	540	430
|der Richter| |der Klägerin| |die Mail |	DAT	590	600	390	470	660	405
	GEN	350	720	420	470	660	405

Pause duration manipulation in ms of the dative sequence tested in perception Experiment 6 (Chapter 4)

sent pair	manipulated pause	condition	original duration	changed to
1.	|P1|	DAT	0.13	0.02
	|P1|	GEN	0.02	0.02
2.	|P1|	DAT	0.04	0.03
	|P1|	GEN	0.03	0.03
3.	|P1|	DAT	0.10	0.04
	|P1|	GEN	0.04	0.04
4.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
5.	|P1|	DAT	0.08	0.03
	|P1|	GEN	0.03	0.03
6.	|P1|	DAT	0.07	0.03
	|P1|	GEN	0.03	0.03
7.	|P1|	DAT	0.09	0.02
	|P1|	GEN	0.02	0.02
8.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
9.	|P1|	DAT	0.06	0.02
	|P1|	GEN	0.02	0.02
10.	|P1|	DAT	0.10	0.02
	|P1|	GEN	0.02	0.02
11.	|P1|	DAT	0.10	0.10
	|P1|	GEN	0.10	0.10
12.	|P1|	DAT	0.10	0.03
	|P1|	GEN	0.03	0.03
13.	|P1|	DAT	0.12	0.03
	|P1|	GEN	0.03	0.03
14.	|P1|	DAT	0.07	0.00
	|P1|	GEN	0.00	0.00
15.	|P1|	DAT	0.08	0.03
	|P1|	GEN	0.03	0.03
16.	|P1|	DAT	0.12	0.03
	|P1|	GEN	0.03	0.03
17.	|P1|	DAT	0.09	0.03
	|P1|	GEN	0.03	0.03
18.	|P1|	DAT	0.09	0.03
	|P1|	GEN	0.03	0.03
19.	|P1|	DAT	0.40	0.01
	|P1|	GEN	0.01	0.01
20.	|P1|	DAT	0.12	0.02
	|P1|	GEN	0.02	0.02

Pause duration manipulation in ms of the experimental conditions tested in Experiment 3 (Chapter 5)

manipulated pause | # |	con	original duration	compressed / expanded to
|der Anwalt | # |der Klägerin|	DAT-nat	388	000
	GEN-nat	000	388
	DAT-con	147	054
	GEN-fall	054	147
|der Bruder| # |der Oberin|	DAT-nat	340	001
	GEN-nat	001	340
	DAT-con	272	039
	GEN-fall	039	272
|der Mörder| # |der Wärterin|	DAT-nat	187	021
	GEN-nat	021	187
	DAT-con	181	040
	GEN-fall	040	181
|der Vater| # |der Schwimmerin|	DAT-nat	232	011
	GEN-nat	011	232
	DAT-con	295	035
	GEN-fall	035	295
|der Zahnarzt| # |der Schwägerin|	DAT-nat	314	000
	GEN-nat	000	314
	DAT-con	291	029
	GEN-fall	029	291
|der Partner| # |der Fahrerin|	DAT-nat	297	000
	GEN-nat	000	297
	DAT-con	550	028
	GEN-fall	028	550
|der Fahrer| # |der Anwältin|	DAT-nat	414	013
	GEN-nat	013	414
	DAT-con	339	042
	GEN-fall	042	339
|der Gönner| # |der Künstlerin|	DAT-nat	389	019
	GEN-nat	019	389
	DAT-con	214	033
	GEN-fall	033	214
|der Bote| # |der Herzogin|	DAT-nat	536	022
	GEN-nat	022	536
	DAT-con	401	032
	GEN-fall	032	401
|der Vogel| # |der Pflegerin|	DAT-nat	208	000
	GEN-nat	000	208
	DAT-con	194	031
	GEN-fall	031	194
|der Trainer| # |der Sportlerin|	DAT-nat	229	000
	GEN-nat	000	229

manipulated pause | # |	con	original duration	compressed / expanded to
	DAT-con	262	038
	GEN-fall	038	262
|der Lehrer| # |der Schülerin|	DAT-nat	338	012
	GEN-nat	012	338
	DAT-con	468	025
	GEN-fall	025	468
|der Kunde| # |der Händlerin|	DAT-nat	278	000
	GEN-nat	000	278
	DAT-con	470	038
	GEN-fall	038	470
|der Lehrling| # |der Gärtnerin|	DAT-nat	108	008
	GEN-nat	008	108
	DAT-con	489	021
	GEN-fall	021	489
|der Junge| # |der Nachbarin|	DAT-nat	319	013
	GEN-nat	013	319
	DAT-con	605	035
	GEN-fall	035	605
|der Zeuge| # |der Sängerin|	DAT-nat	460	020
	GEN-nat	020	460
	DAT-con	619	034
	GEN-fall	034	619

Segmental duration manipulation of NP1 in ms of the conditions tested in Experiment 4 (Chapter 5)

manipulated words	con	original duration	compressed / expanded to
|der Anwalt |	DAT-nat	594	511
	GEN-nat	511	594
	DAT-con	639	616
	GEN-fall	616	639
|der Bruder|	DAT-nat	624	465
	GEN-nat	465	624
	DAT-con	619	565
	GEN-fall	565	619
|der Mörder|	DAT-nat	541	479
	GEN-nat	479	541
	DAT-con	590	573
	GEN-fall	573	590
|der Vater|	DAT-nat	588	478
	GEN-nat	478	588
	DAT-con	617	507
	GEN-fall	507	617
|der Zahnarzt|	DAT-nat	825	622
	GEN-nat	622	825
	DAT-con	838	690
	GEN-fall	690	838
|der Partner|	DAT-nat	644	578
	GEN-nat	578	644
	DAT-con	677	660
	GEN-fall	660	677
|der Fahrer|	DAT-nat	643	450
	GEN-nat	450	643
	DAT-con	608	492
	GEN-fall	492	608
|der Gönner|	DAT-nat	501	409
	GEN-nat	409	501
	DAT-con	519	432
	GEN-fall	432	519
|der Bote|	DAT-nat	550	421
	GEN-nat	421	550
	DAT-con	563	499
	GEN-fall	499	563
|der Vogel|	DAT-nat	597	555
	GEN-nat	555	597
	DAT-con	554	491
	GEN-fall	491	554
|der Trainer|	DAT-nat	666	546
	GEN-nat	546	666

manipulated words	con	original duration	compressed / expanded to
	DAT-con	639	559
	GEN-fall	559	639
|der Lehrer|	DAT-nat	629	525
	GEN-nat	525	629
	DAT-con	605	468
	GEN-fall	468	605
|der Kunde|	DAT-nat	659	554
	GEN-nat	554	659
	DAT-con	683	518
	GEN-fall	518	683
|der Lehrling|	DAT-nat	613	533
	GEN-nat	533	613
	DAT-con	654	571
	GEN-fall	571	654
|der Junge|	DAT-nat	545	455
	GEN-nat	455	545
	DAT-con	547	427
	GEN-fall	427	547
|der Zeuge|	DAT-nat	679	594
	GEN-nat	594	679
	DAT-con	679	621
	GEN-fall	621	679

Upstep and downstep manipulation on NP2 of the experimental conditions tested in Experiment 5 (Chapter 5)

manipulated words	con	original f0-max (Hz)	increased / reduced to
|der Klägerin |	DAT-nat	165	139
	GEN-nat	139	165
|der Oberin|	DAT-nat	151	141
	GEN-nat	141	151
|der Wärterin|	DAT-nat	134	131
	GEN-nat	131	134
|der Schwimmerin|	DAT-nat	160	135
	GEN-nat	135	160
|der Schwägerin|	DAT-nat	141	130
	GEN-nat	130	141
|der Fahrerin|	DAT-nat	151	139
	GEN-nat	139	151
|der Anwältin|	DAT-nat	153	138
	GEN-nat	138	153
|der Künstlerin|	DAT-nat	154	140
	GEN-nat	140	154
|der Herzogin|	DAT-nat	166	134
	GEN-nat	134	166
|der Pflegerin|	DAT-nat	149	142
	GEN-nat	142	149
|der Sportlerin|	DAT-nat	172	133
	GEN-nat	133	172
|der Schülerin|	DAT-nat	172	155
	GEN-nat	155	172
|der Händlerin|	DAT-nat	175	143
	GEN-nat	143	175
|der Gärtnerin|	DAT-nat	167	148
	GEN-nat	148	167
|der Nachbarin|	DAT-nat	168	146
	GEN-nat	146	168
|der Sängerin|	DAT-nat	165	141
	GEN-nat	141	165

Hiermit erkläre ich, dass ich die vorliegende Dissertation *Prosodic cue weighting in sentence comprehension: Processing German case ambiguous structures* selbständig und nur unter Zuhilfenahme der angegebenen Quellen und Literatur angefertigt habe. Ferner erkläre ich, dass die Arbeit noch nicht in einem anderen Studiengang als Prüfungsleistung verwendet wurde.

Potsdam, den 14. November 2013

Anja Gollrad