Linguistic Rhythm and Sentence Comprehension in Reading

Inauguraldissertation

zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Neuere Philologien
der Johann Wolfgang Goethe-Universität
zu Frankfurt am Main

vorgelegt von Gerrit Kentner
aus Oldenburg (i.O.)

2010
(Einreichungsjahr)

2012
(Erscheinungsjahr)

Gutachter:

Prof. Dr. Caroline Féry (Betreuerin)

Prof. Dr. Shravan Vasishth

Tag der Disputation: 14. Februar 2011

Für Jenny und Tilo
und Olaf

Danksagung

Die vorliegende Arbeit ist zu einem großen Teil im Rahmen des Projekts “Prosody in Parsing” im DFG Schwerpunktprogramm “Sprachlautliche Kompetenz” entstanden. Ich danke Caroline Féry, Shravan Vasishth und Frank Kügler, die die Gelder eingeworben und mir damit ein Dissertationsprojekt ermöglicht haben, welches ich in der außerordentlich inspirierenden Atmosphäre des Potsdamer Instituts für Linguistik angehen konnte. Allen Dreien gebührt darüber hinaus Dank für vielfache Ermutigung, kollegiale Freundschaft und maßgebliche Impulse für meine akademische Ausbildung:
Caroline hat mir mit der Potsdamer und der Frankfurter Stelle den Freiraum geschaffen, der für die Ideenentwicklung zu dieser Dissertation unabdingbar ist. Ihr verdanke ich wertvolle inhaltliche Anregungen für diese Arbeit und angedachte Folgeprojekte. Ich freue mich sehr, weiterhin mit ihr zusammen zu arbeiten.
Without Shravan’s constant and spirited support, this thesis would not have the theoretical and empirical rigor the subject deserves. I am most grateful for his introduction into the methodology of eye tracking research, for providing the lab, and especially for teaching me fearlessness against the statistical intricacies that one comes across in experimental research.
Frank danke ich für seine phonetische Expertise, zahlreiche Anregungen und für seine Freundschaft. Insbesondere seinem Ansporn verdankt das PiP-Projekt, dem ich mich weiterhin verbunden fühle, eine zweite Lebensphase.

Die Idee zu dieser Arbeit ist eng verbunden mit dem Seminar “Rhythmus und Syntax”, das Ralf Vogel und Ruben van de Vijver im Sommer 2003 angeboten haben. Ihnen danke ich für diese weit über das Jahr 2003 hinausgehende Inspiration und für die Gelegenheit, meine Arbeit bei der 2009er DGfS-Tagung in Osnabrück vorzustellen.
Sincere thanks go to Lyn Frazier, who showed her interest in this work and invested a lot of time for fruitful discussions. Auch Markus Bader sei für sein Interesse und seine hilfreichen Kommentare gedankt. Audiences at “Experimental and Theoretical Advances in Prosody” (Cornell, 2008), the “Potsdam–Lund Prosody Meeting” (Potsdam, 2008) and ”Experimental Studies on Intonation” (Potsdam, 2009) provided encouraging comments and helpful suggestions.

Mit meinen Potsdamer Büro- und Projektkollegen hatte ich wundervolle gemeinsame Jahre in Raum 002 in Haus 35. Anja Gollrad teilt mein Interesse an der Prosodie und ich danke ihr für die vielen Diskussionen und freundschaftlichen Gespräche, die weit über das Akademische hinaus gingen.
I am very lucky to have become friends with Umesh Patil, who has been very

supportive and interested in matters of my work and beyond. The discussions with him were extremely helpful and always fun. I admire the crafty style of his badminton play. I also thank Stella Gryllia for her collegiality and for the fruitful and friendly discussions we had.

Pavel Logačev danke ich für seine immens wichtige Hilfe bei der Verarbeitung der eye-tracking-Daten. Großer Dank geht an Timo Baumann, der mich geduldig in die Geheimnisse der automatischen Annotierung eingeführt hat.

Ich danke allen, die außerdem zu dem stimulierenden Umfeld beigetragen haben, in dem ich mich in Golm bewegen durfte. Ohne Anspruch auf Vollständigkeit nenne ich Anja Arnhold, Dinah Baer-Henney, Jutta Boethke, Gerlof Bouma, Okko Buss, Heiner Drenhaus, Felix Engelmann, Annett Eßlinger, Gisbert Fanselow, Tom Fritzsche, Susanne Genzel, Peter beim Graben, Judith Heide, Jule Holzgrefe, Robin Hörnig, Shin Ishihara, Titus von der Malsburg, Ines Mauer, Daniel Schad, Fabian Schubö, Rukshin Shaher, Esther Sommerfeld, Hans Trukenbrod, Hubert Truckenbrodt, Isabell Wartenburger, Thomas Weskott, Thomas Wiederhold und Malte Zimmermann.

Ich danke auch den Frankfurter Kollegen, die mich hier freundlich aufgenommen haben: Sascha Dümig, Eric Fuß, Anja Geumann, Günther Grewendorf, Angela Grimm, Daniel Gutzmann, Fatima Hamlaoui, Katharina Hartmann, Cécile Meier, Sara Myrberg, Birgit Nutz, Christian Plunze, Henning Reetz, Maren Schmidt-Kassow, Viola Schmitt, Matthias Schulze-Bünte, Anke Stakemann, Anna Volodina, Frank Zimmerer und Ede Zimmermann.
Ich freue mich dass Andreas Pankau und Shin Ishihara als ebenfalls Ex-Potsdamer Kollegen am Frankfurter Institut sind (letzterer sogar als Büronachbar).

Schließlich danke ich meinen Eltern und Geschwistern von Herzen für ihre bedingungslose Unterstützung und Ermutigung.
Die Arbeit an dieser Dissertation hat mir Freude gemacht; es ist aber gut, dass es im Leben noch größere Freuden gibt. Ich widme diese Arbeit Jenny und Tilo.

Nachtrag (Oktober 2012): Seit ich diese Dissertation im Oktober 2010 eingereicht habe, sind zwei Jahre vergangen. Zwei Kapitel (2 und 3) sind seitdem mit einigen Änderungen zur Veröffentlichung angenommen worden; das 4. Kapitel soll in einem Sonderband von The Linguistic Review erscheinen. Ich danke den Herausgebern und den (anonymen) Gutachtern, die die Veröffentlichung mit wohlwollend-kritischen Kommentaren begleitet haben; sie haben damit wesentlich zur Verbesserung der Arbeit beigetragen. Jenny Kentner und insbesondere Kirsten Brock danke ich für das sorgfältige Lektorat.

Contents

Chapter 1

Introduction

Writing never really got around to providing a regular way of marking accent [...], and it has virtually disregarded rhythm and intonation.
Dwight Bolinger

... there has never been a line read that I didn't hear. As my eyes followed the sentence, a voice was saying it silently to me.
Eudora Welty

1.1 General scope and aims

This dissertation is concerned with the role of prosody and, specifically, linguistic rhythm for the syntactic processing of written text. My aim is to put forward, provide evidence for, and defend the following claims:

- While processing written sentences, readers make use of their phonological knowledge and generate a mental prosodic-phonological representation of the printed text.

- The mental prosodic representation is constructed in accordance with a syntactic description of the written string. Constraints at the interface of syntax and phonology provide for the compatibility of the syntactic analysis and the (mental) prosodic rendition of the sentence.

- The implicit prosodic structure readers impose on the written string entails phonological phrasing and accentuation, but also lower level prosodic features such as linguistic rhythm which emerges from the pattern of stressed and unstressed syllables.

- Phonological well-formedness conditions accompany and influence the process of syntactic parsing in reading from the very beginning, i.e. already at the level of recognizing lexical categories. At points of underspecified syntactic structure, syntactic parsing decisions may be made on the basis of phonological constraints alone.

- In reading, the implicit local lexical-prosodic information may be more readily available to the processing mechanism than higher-level discourse structural representations and consequently may have more immediate influence on sentence processing.

- The process of sentence comprehension in reading is conditioned by factors that are geared towards sentence production.

- The interplay of syntactic and phonological processes in reading can be explained with recourse to a performance-compatible competence grammar.

Justification for the first claim comes from readers' introspection: most literate people are familiar with the experience of subvocalization or the phenomenon of 'inner speech' in reading, a form of silent articulation and prosodification that not only comprises phonological structure but may also include paralinguistic features such as voice quality (Chafe, 1988). Despite being inaudible, it has sometimes been characterized as being essentially speechlike (e.g. Ashby and Martin, 2008). As natural as this characterization may seem, it prompts the question of how the apparently rich prosodic representation is generated given that written text usually lacks explicit cues to prosody. Clearly, it has to be produced on the basis of more than just the printed letter sequence.

The other hypotheses require a good deal of careful examination. The evidence in favor of these hypotheses is hardly traceable on the basis of introspection alone. More importantly, these claims have far reaching implications for the role of phonology in parsing and, more generally, for the conception of the grammar the sentence comprehension mechanism makes use of. Thus, although this dissertation concentrates on the examination of prosodic effects in one particular mode of language processing, viz. reading, it has something to say about the general role of prosody in relation to other aspects of grammar as well.

Beyond their psycholinguistic importance, these points suggest a model of grammatical competence in which constraints from various domains (syntax, semantics, pragmatics, discourse structure, and phonology) interact in providing the possible structural, i.e. grammatical descriptions. Applying this constraint satisfaction approach to language allows us to readily relate the grammar to processing phenomena. At the same time, it provides an architecture that captures comprehension and production processes – both are engaged in reading – within a single framework. The experiments presented in this dissertation thus inspire the formulation of a performance model that looks directly into the competence grammar to derive parsing predictions. I concur with Jackendoff (2003) in stating that such an integrative approach to linguistic theory and psycholinguistic evidence 'ought to be favored over one that creates hard boundaries between competence and performance.' (Jackendoff, 2003, p. 197).

1.2 The architecture of this dissertation

Rather than a comprehensive monograph, this dissertation is arranged as a collection of thematically closely related but fairly independent studies. The reading studies presented here are designed to provide data on the role of linguistic rhythm in oral reading and written sentence comprehension, culminating in the elaboration of a sentence processing model that is designed to capture the role of prosodic factors in reading.

The remainder of this first chapter introduces some background on supralexical prosody with its relation to syntax, and sets the stage for the case under scrutiny: the role of linguistic rhythm for sentence comprehension in reading. The following two chapters report on experiments that are designed to study the reading performance on exemplary sentence structures, which may shed light on the function of linguistic rhythm in written sentence comprehension:

Chapter 2 presents an experimental study on the effects of linguistic rhythm on processing non-canonical structures in oral reading, examining the interplay of syntax, focus structure, prosody, and working memory. The experiment attests that a written word sequence that forces readers to deviate from the optimal rhythmic alternation of stressed and unstressed syllables may critically hamper the processing of complex, non-canonical sentences. At points of high memory demand, the cognitive load associated with a stress clash configuration may lead to the (temporary) overwriting of memory traces that are necessary for the processing of syntactic long distance dependencies. I argue that this overwriting becomes possible because the clash information, which is implicitly coded in the written string, is processed immediately while access to stored information in working memory may be tardy.

Chapter 3 affords more data on effects of linguistic rhythm on written sentence comprehension. Two reading experiments examine the influence of stress-based linguistic rhythm on the resolution of local lexical-syntactic ambiguities. Both speech production data from unprepared oral reading as well as eye-tracking results from silent reading demonstrate that readers favor syntactic analyses that allow for a prosodic representation in which stressed and unstressed syllables alternate rhythmically. The findings contribute evidence confirming immediate and guiding effects of linguistic rhythm on the earliest stages of syntactic parsing in reading.

In **Chapter 4**, the performance data from the experiment reported in Chapter 3 are modeled as an incremental constraint satisfaction process in the framework of an Optimality Theoretic parsing account. Solely making use of constraints derived from competence grammar, the model is capable of capturing the data and advocates the simultaneous application of syntactic, prosodic and syntax-phonology interface constraints in incremental processing. The model predicts that, in the case of syntactic indetermination, weak prosodic constraints may decide about syntactic ambiguity resolution. The performance-compatible OT grammar integrates the processes of syntactic parsing and prosodification in reading, hence dissolving the strict separation of language production and comprehension. At the same time the OT model endorses a bidirectional relationship between syntax and phonology in grammar.

Chapters 2, 3, and 4 each provide specific background and discussion so that they can be read independently. This also means that the reader will come across an hopefully tolerable amount of redundancy.

In **Chapter 5**, the dissertation concludes with a summary.

1.3 Background

1.3.1 Supralexical prosody

Prosody is commonly understood as comprising the suprasegmental aspects of the speech signal. Prosodic features can be observed phonetically in terms of fundamental frequency, intensity, and duration; their respective physical parameters can be objectively measured; such absolute measurements, however, are not in a one-to-one relationship with the psychological percept they might induce and are, consequently, only loosely related to the grammatical system of a language. The vocabulary for the linguistic description of prosodic phenomena such as pitch and duration is mostly confined to relational notions such as 'higher than' or 'longer than'. This is not to be understood as a shortcoming of the human perceptual system but reflects the fact that a given prosodic phenomenon is generally interpreted relative to the surrounding prosodic events. The linguistically relevant aspects of prosody have been captured within a hierarchical system that implicitly makes use of the relational notions just alluded to: The prosodic hierarchy (Nespor and Vogel, 2007; Selkirk, 1984) defines prosodic constituents of various levels that are organized according to their respective size and prominence: the smaller prosodic constituents, e.g. syllables, are embedded in larger ones such as words, prosodic phrases or utterances. The relative prominence of the different levels corresponds with their respective size and is normally expressed by stress on (lexically) designated syllables within the prosodic level under consideration. There is, correspondingly, no invariant phonetic correlate of stress. Rather, the realization of stress on a given syllable heavily depends on the nature of the domain for which the syllable serves as prosodic a head. Generally, stressed syllables are marked by longer duration, more pitch deflection and higher intensity compared to unstressed syllables; stressed syllables that function as a head of phonological phrases or intonational phrases are generally pitch accented, i.e. marked by a clear and audible deflection of pitch and longer duration as, for example, in the first syllable of the last word in (1). I follow the convention to mark pitch accented words with small caps and the stressed syllable (the one that potentially carries accent) with underlines.

(1) He went <u>SHOP</u>PING.

The prosodic hierarchy (2) explicates the organization of prosodic constituents. The levels within this hierarchy are defined on the basis of (language specific) phonological processes for which the constituents provide a domain. The higher domains also have relatively clear phonetic correlates. As an example, the intonational phrase (IP) is demarcated by boundary tones (also realized as deflection of pitch), lengthening of the final syllable(s) and, optionally, pauses at their end.

(2) The prosodic hierarchy (following Nespor & Vogel 1986, 2007; Selkirk 1984)

Utterance (υ)
|
Intonational Phrase (IP)
|
Phonological Phrase (φ)
|
Prosodic Word (ω)
|
Foot (Ft)
|
Syllable (σ)

It is important to note that certain levels of the prosodic hierarchy are closely related to, albeit not necessarily isomorphic with, non-prosodic entities: phonological words often coincide with syntactic words while the prosodic phrase roughly corresponds to the maximal projection of a lexical head. A full sentence normally correlates with an IP, although it might contain more than one IP if its length or syntactic complexity exceed a certain threshold. In such a case, the distribution of IPs normally coincides with the distribution of larger syntactic domains. Consider the phrasings in (3) (example following Mark Liberman, cited in Selkirk, 1995, IP boundaries marked with '//'). The sentence may be split in two IPs with the subject wrapped in one IP and the verb phrase in the other (3-a), yielding a perfect match between syntactic and prosodic constituents. Alternatively, the prepositional phrase may be uttered as a parenthetical or afterthought. In this case, the sentence is distributed over three IPs as in (3-b). A sole IP boundary splitting up the subject and phrasing the PP together with the VP as in (3-c), however, is illicit. Apparently, as the PP and VP do not form a syntactic constituent, IP boundaries are to be set in accord with syntactic structure.

(3) a. Three mathematicians in ten // derive a lemma.
b. Three mathematicians // in ten // derive a lemma.
c. *Three mathematicians // in ten derive a lemma.

It is, by hypothesis, the correspondence between prosodic and syntactic constituents that makes the syntactic analysis of oral language fast and seemingly effortless. However, prosodic structure does not reflect syntactic structure unambiguously. Under certain conditions, prosodic boundaries may interrupt rather than demarcate syntactic phrases. Consider, for example, the following prosodic phrasing in Italian (taken from Ghini, 1993, phonological phrases are marked with round, and syntactic constituents with square brackets).

(4) (Prendera [granchi)$_{\varphi}$ (di ogni specie)$_{\varphi}$]$_{DP}$
will catch crabs of every kind
'He will catch every kind of crab'

In (4), the object DP is interrupted by a phonological phrase boundary. A perfect match between syntactic structure and prosodic rendition would require that the object DP be wrapped within a single phonological phrase. The mismatch, according to Ghini (1993), is due to a requirement on prosodic structure that demands the formation of roughly balanced, i.e. equal-sized phonological phrases. This requirement would be violated if the sentence was chunked into the short verb on the one hand and the complex object on the other. That is, syntax-prosody correspondence requirements interact with constraints on the well-formedness of prosodic structure. In the variety of Italian studied by Ghini (1993), the prosodic balance requirement apparently overrides the syntax-prosody alignment constraint.

Stress and linguistic rhythm

The example from Italian already indicates that prosody has its independent existence with own regularities at all levels of the prosodic hierarchy. A further example of a prosodic regularity, one that will be relevant throughout this dissertation, concerns the concatenation of syllables and the prominence pattern that emerges consequently. Within a phonological phrase, the sequence of syllables is preferably rhythmic in the sense that stressed and unstressed syllables alternate. The adjacency of stressed syllables (stress clash) is generally avoided and so are longer sequences of unstressed syllables, so-called lapses (Hayes, 1995; Kager, 1989; Liberman and Prince, 1977; Selkirk, 1984).

One piece of evidence for stress clash avoidance comes from the distribution of mono- and disyllabic variants of participle forms in English. Schlüter (2003) offers an analysis of the synonymous and etymologically related English participle forms 'lit' and 'lighted' on the basis of two large scale newspaper corpora. Both variants may appear as predicative or attributive adjective. Schlüter finds that 'lit' has the clear majority (94% of all tokens), while the form 'lighted' is consistently preferred in the single attributive use of the participle, as for example in *'the puddle of lighted* PETROL*'*. The reason for the preference of the otherwise marginal 'lighted', according to Schlüter (2003), lies in the fact that the suffix *-ed* provides an additional unstressed syllable that prohibits a stress clash in this position. This is evident if the alternative form *'lit* PETROL*'* is considered. Given that the preponderance of nouns in English feature initial stress, clashes are unavoidable in the case of monosyllabic attributive adjectives. Accordingly, if speakers (or writers, for that matter) have the choice, they might resort to the disyllabic variant.

More evidence for the markedness of stress clash comes from instances of stress shift, i.e. the deviance from the citation form stress pattern triggered by a clash of adjacent stressed syllables as in (5) (example following Kiparsky, 1966). The verb *mitnehmen* bears main stress on the initial syllable in citation form (stressed syllables underlined, phrasal stress marked by small caps). In connection with a monosyllabic object (which generally receives main phrase stress, cf. Truckenbrodt, 2006) a stress clash results. The clash is normally resolved by shifting the main stress of the verb to the second syllable, thus achieving rhythmic alternation.

(5) a. mitnehmen *(citation form)*
take along
b. #das BUCH mitnehmen
b'. das BUCH mitnehmen
take along the book

Both the findings by Schlüter (2003) and stress shift cases like (5) show that the preference for rhythmic alternation operates beyond word boundaries. Correspondingly, the avoidance of stress clash is a *supralexical* phonotactic constraint. Due to its supralexical nature, it may have a bearing not only on lexical form but, by implication, also on syntactic form. Consider, in this regard, a case in which speakers have the choice between two (semantically roughly equivalent) syntactic constructions as in (6) (example following Speyer, 2008, 2010). Given the question in (6) one could answer with a regular SVO sentence (6-a). Alternatively, one might want to front, or topicalize the object *beans*. This is a licit way to utter the respective proposition since legumes are established as the discourse topic in the question. Topicalization, however, is preferred only under the condition that the subject, which appears in second position now, be realized as a weak pronoun (6-b) thus warranting rhythmic alternation. Realizing the topicalized object directly preceding a full lexical subject creates a stress clash, rendering (6-c) infelicitous unless a clear pause be inserted between the topicalized object and the subject (Speyer, 2008, 2010).

(6) What kind of legumes does John like?
a. JOHN likes BEANS.
b. BEANS he LIKES.
c. #BEANS JOHN likes.

The example suggests that the choice of the syntactic form is not independent from the prosodic rendition. The data in (6) justify the conclusion that, ceteris paribus, word orders that allow for rhythmic alternation are favored over word orders that induce stress clashes.

A fully fledged grammar that claims to cover syntax-prosody interactions of the kind reviewed here, therefore, has to deploy purely prosodic regularities along with syntax-prosody interface conditions.

Before conceptualizing the proposed interaction of rhythm and syntax, I will briefly discuss the importance of prosody in auditory sentence comprehension and provide background on the role of prosody in understanding written language.

Prosody in auditory sentence comprehension

In auditory language processing, listeners are provided with, and have been shown to use, various prosodic cues that potentially point the reader to syntactic structure of the speech input (cf. Cutler et al., 1997, for a review). Prosodic features that have been shown to influence syntactic parsing and sentence interpretation are prosodic phrase boundaries (Marslen-Wilson et al., 1992; Pynte and Prieur, 1996; Kjelgaard and Speer, 1999; Frazier et al., 2006, among others), accentuation (e.g. Carlson, 2001; Ito and Speer, 2008; Schafer et al., 2000; Steinhauer et al., 1999), the size of constituents in terms of syllables or prosodic words (e.g. Augurzky, 2006; Fodor, 1998) and, to a lesser degree, the distribution of stressed and unstressed syllables (Warren et al., 1995). All these prosodic characteristics are an integral part of sentence phonology and are often conditioned by the syntactic environment. Prosody, however, has various functions aside from reflecting syntactic structure; among other things, it may signal discourse structure and also paralinguistic phenomena such as attentional or emotional states. The distinctness of the translation between syntactic facts and prosodic structure heavily depends on the prosodic context: independent factors such as speech tempo and utterance length condition the clarity of the prosodic rendition of syntactic constituent boundaries. Moreover, sentence prosody is variable and, consequently, not completely predictable from the syntactic form. Any given structure may receive a range of valid alternative prosodic realizations. Listeners seem to have a tacit knowledge about the uncertain relationship of prosodic and syntactic structure; correspondingly, not all prosodic information that may potentially guide the listener in sentence comprehension is actually used. As Watt and Murray (1996) put it:

> *There exists a subset of 'natural' prosodies which, although intelligibly different and capable of influencing appropriateness judgments, exert no influence on initial parsing processes.*

According to this view, sentence prosody may reflect syntactic structure but the listener does not necessarily make immediate use of this information.

Other authors, however, reinforce the importance of prosody in language comprehension emphasizing its predictive value. Experimental evidence suggests the ability of listeners to anticipate the amount of as yet unheard prosodic material on the basis of local prosodic cues. Grosjean (1983) as well as Grosjean and Hirt

(1996) found that native English listeners could reliably predict the number of words to the end of the sentence on the basis of prosodic cues of a sentence fragment – arguably an important cue for the syntactic analysis. Similarly, Snedeker and Trueswell (2003) report that prosodic cues in the input may bias listeners towards one specific reading of a syntactically ambiguous sentence even before the ambiguous region is reached.

What, then, is the role of prosody in sentence comprehension? Given the equivocal results, it should be clear that accounts which assign the interpretation of prosody a fixed place within a linear processing chain of language comprehension run into problems. A more viable alternative is one in which the processing mechanism evaluates prosodic, syntactic and other information simultaneously as it becomes available in the input. Thus, the influence of a certain prosodic feature crucially depends on the strength of concurring input features. On this view, prosodic features are evaluated not before or after, but relative to, the simultaneously available linguistic (and even non-linguistic) information. That is, the effect a prosodic phenomenon has on sentence comprehension is not solely determined by the clarity of its phonetic expression but also by the strength of the ontologically different concurring cues to meaning such as word order, morphosyntactic information, and discourse context.

1.3.2 Phonology in reading

While the prevailing share of our communication in everyday life is oral, written communication does occupy large parts of our communicative actions. Reading is certainly a secondary form of language perception, the mastery of which normally requires competence in the oral modality. Despite the ontogenetic and phylogenetic dependence of written from oral language, most of our reading is silent rather than oral. Given the obvious soundlessness of silent reading (grunts and smacking aside), the following question arises: what role does the readers' knowledge of language sounds, i.e. the phonology, play in processing the written modality? Let's ask a man of letters, the late philosopher Friedrich Nietzsche:

> *The German does not read aloud, does not read for the ear but merely with his eyes: he has put his ears away in the drawer.*
>
> Friedrich Nietzsche, Beyond Good and Evil, §247

The 'ear' in Nietzsche's quote can of course be understood as a metaphor for the phonological system (i.e. the system dealing with the sounds of language)[1]. In this

[1]Nietzsche probably didn't intend to place on the market strong predictions about the role of phonology in language processing. The context of the quote suggests that he is interested in the cross-cultural relevance of, and handling with, written texts.

sense, Nietzsche would argue for the strong claim that readers (or the German reader for that matter) use a non-phonological route from print to meaning.

It turns out that there is some support for this view from the neurolinguistic domain. Indeed, there is evidence to show that, at least under pathological conditions such as 'deep dyslexia' or 'phonemic dyslexia', reading comprehension may be spared even if phonological recoding is heavily impaired (e.g. Beauvois and Dérouesné, 1979; Saffran and Marin, 1977; Shallice and Warrington, 1975); these studies suggest that patients with certain neurological disorders may access word meanings directly from the printed word bypassing phonology. If patients show this skill, healthy readers should certainly have this capacity, too. The evidence for reading comprehension without phonological recoding, however, seems to be limited to the word level and it is questionable whether it can be generalized to the sentence level given that the patients described also suffer from deficits that are attributable to syntactic processing.

More importantly, based on a comprehensive research survey on written word recognition in healthy populations, Frost (1998) concludes that non-phonological reading is the exception rather than the rule. Many of the studies Frost reviewed do find unequivocal and immediate effects of phonological recoding on word recognition in reading.

Beyond the word level, there is, to the best of my knowledge, no evidence that would suggest that readers arrive at meaning without recourse to phonology or at least some kind of sublexical representation. Introspection suggests that sentence reading involves the mental construction of sound images that include not only the sentence's prosodic-phonological structure but may even comprise paralinguistic features such as speech tempo or voice quality (Chafe, 1988). This by itself does not warrant the claim that subvocalization or phonological recoding is necessary for sentence comprehension. However, artificially disrupting the subvocalization process seems to hamper sentence comprehension (e.g. Kleiman, 1975; Slowiaczek et al., 1980). Kleiman (1975) proposes that phonological structure is generated during silent reading in order to provide a format necessary for the storage of linguistic information in working memory. This in turn, the argument goes, is obligatorily required for sentence comprehension.

Implicit prosody

More recent sentence processing research indicates that certain features of the prosody that naturally emerges during the subvocalization process (also called implicit or silent prosody) influences the resolution of syntactic ambiguities. Bader (1996b, 1998) was the first to argue that implicit prosody partly determines the difficulty of reanalyzing temporarily ambiguous sentences in reading. If two competing syntactic structures differ with respect to their prosodic renditions, reanalysis, if necessary, is predicted to be more difficult compared to cases in which the prosodic structures of the concurring syntactic interpretations coincide. This hypothesis may help to explain why the recovery from a garden path sentence like (7-a) feels more difficult than a reanalysis that might be required for parsing (8-b). Note that both structures are syntactically similar in that readers may interpret the ambiguous item (*the door* in (7) and *the boy* in (8)) either as a direct object of the first clause or as the subject of the following clause. For the oral rendition of the sentences in (7), a clear intonational phrase break is required in either version to separate matrix clause and embedded clause. The two readings of (7) are thus distinguished by the position of the prosodic phrase break. According to Bader's Prosodic Constraint on Reanalysis, the unprepared reader commits himself to the preferred syntactic structure and the concomitant prosody. If the chosen syntactic analysis turns out to be incompatible with the disambiguating information, both syntactic structure and prosody have to be reanalyzed. The additional prosodic reanalysis is assumed to be cognitively costly. As for (8), no prosodic feature clearly distinguishes the two readings. Syntactic reanalysis can therefore proceed swiftly without concomitant prosodic reanalysis.

(7) a. Whenever the guard checks the door // it's locked
    b. Whenever the guard checks // the door is locked

example following Speer et al. (1996)

(8) a. Mary knows the boy on the bench.
    b. Mary knows the boy is sleeping.

example taken from Wagner and Watson (2010)

Observe that an interpretation of this type gives temporal and causal priority to the syntactic analysis of the written string over its prosodic interpretation. According to Bader (1998), implicit prosody only affects the revision of the initial syntactic analysis but cannot guide the early stages of syntactic parsing itself. This view is reminiscent of the stance of Watt and Murray (1996) on the role of prosody in auditory sentence comprehension, stating that initial parsing stages are blind to (at least certain) prosodic information.

Fodor (2002) proposes a more balanced influence of prosody and syntax in suggesting that the prosodic interpretation and the syntactic analysis mutually interact in written sentence comprehension. Her view conforms to an idea that Bierwisch (1966) put forward in the context of oral language processing:

> *Es ist sehr wohl möglich, dass ein Sprecher sich durch ein begonnenes Intonationsmuster zur Wahl einer bestimmten syntaktischen Struktur veranlasst fühlt* [...].
> It is very well possible that a speaker feels himself prompted to choose a

> certain syntactic structure due to the intonation pattern that he has generated.
>
> Bierwisch (1966), p. 105.; translation: G.K.

If transferred to reading (in which the choice of syntactic structure is heavily constrained by the fixed word order) an interesting idea arises: the prosody that readers (implicitly or explicitly) impose onto the written string may guide their syntactic interpretation of the word sequence. Fodor (1998, 2002) subscribes to this idea. In her Implicit Prosody Hypothesis, she states:

> *'In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction.'*

This means that the way a written sentence is prosodically rendered by an inner voice provides distinguishing information when syntactic and semantic information alone leave the input ambiguous. Fodor ascribes implicit prosody a guiding function in sentence comprehension when she notes that '[a prosodic break] can bias the resolution of a syntactic ambiguity just as a prosodic break in a spoken sentence does' (Fodor, 2002, p. 2).

1.3.3 Research questions and hypotheses

As successful as the implicit prosody idea has proven for reconciling hitherto puzzling cross-linguistic differences concerning clause attachment preferences (e.g. Augurzky, 2006; Fodor, 2002; Hwang and Schafer, 2009; Jun, 2003), two important issues as to the relation of prosody and syntax in reading performance remain open:

First, so far, all studies examining effects of implicit prosody in reading tacitly or explicitly assume that at least a low-level syntactic analysis is necessary in order for implicit prosody to show effects. This view is most clearly stated in Augurzky (2006):

> *[...] the parser initially leaves the prosodic analysis underspecified. Early decisions are thus supposed to be guided by the more reliable structural information alone.*
>
> Augurzky (2006), p. 206.

On this view, however, the strength of the implicit prosody idea is considerably weakened - the guiding function of prosody in sentence comprehension is severely questioned if it is seen as a by-product of some syntactic pre-processing.

Secondly, the considerable inter-speaker variability on the level of what is commonly understood as sentence prosody (intonation, prosodic phrasing, accentuation) has so far impeded the establishment of a firm role for prosody in comprehending written text. Given the range of possible and admissible prosodic realizations for a sentence, the notion of 'the most natural' or 'default prosody' must remain elusive. Indeed, as Swets et al. (2007) find, the large individual differences in working memory capacity may critically affect the way a sentence is prosodically parceled in silent reading. Readers with lower working memory capacity tend to chunk the sentences into smaller prosodic units than readers with higher memory spans. In the light of such findings one may ask whether prosodic effects on syntactic parsing in reading can be interpreted as emerging from prosodic grammar or whether they are better captured with recourse to the notion of memory chunks.

The question, then, is: How immediate is the prosodic influence on sentence processing in reading? What is the time course of prosodic processing relative to syntactic processing and how do syntactic parsing and implicit prosody in reading interact?

To answer these questions with respect to effects of implicit prosody on reading, one ought to focus on those prosodic factors that are i) sufficiently stable across individual speakers to warrant clear statements on their nature, and that ii) promise to have immediate effects on syntactic structure building.

A prosodic feature that potentially meets these requirements is lexical stress and the prominence pattern arising from the concatenation of stressed and unstressed syllables. German speakers have very clear and consistent intuitions about the position of stress in words (e.g. Wiese, 2000), a fact that is supported by the explicit marking of lexical stress in dictionaries. Regarding written text processing, word stress has been identified as one of the earliest lexical information available to the reader, i.e. within the first 100ms on visual encounter of the word (Ashby and Martin, 2008). There is thus good reason to assume that, while we read and parse the sequence of lexical items, a pattern of stressed and unstressed syllables unfolds within our internal hearing (and through the eyes). Given the immediacy of the stress information, this pattern may have an early influence on the analysis of the text.

As discussed above, supralexical phonotactic constraints operate on the sequence of stressed and unstressed syllables. Stress clashes are particularly avoided. Just as the avoidance of clash may bear on the choice of the syntactic construction in language production (cf. (6) above), it may have an influence on the syntactic structure that readers assign to the lexical string in reading and, accordingly, on written text comprehension. I therefore hypothesize that, whenever the written word sequence is underspecified with respect to syntactic structure and syllabic prominence, the reader will choose whatever structure best satisfies the relevant syntactic and prosodic constraints. In the case of an ambiguous or underspecified word sequence, the reader will, everything else being equal, favor a syntactic analysis that allows for the rhythmic alternation of stressed and unstressed syllables.

In this dissertation, I set out to test this hypothesis empirically. In the following chapters I will report several reading experiments designed to gauge the influence of linguistic rhythm on the comprehension of written sentences. The largely consistent findings of the reading experiments give credence to the hypothesis that linguistic rhythm, or more precisely: the expectation of rhythmic alternation, affects the syntactic analysis of written text.

Based on the results, I will propose a model of sentence comprehension in reading that captures the processing facts by incorporating a grammar in which syntactic and prosodic constraints interact together with requirements on the syntax-prosody interface. In this grammatical constraint satisfaction model, the different types of constraints simultaneously contribute to incremental sentence processing. Moreover, the data and the model suggest that written sentence comprehension is best understood as a process that is partly driven by factors that are traditionally ascribed to sentence production.

Chapter 2

Stress clash hampers processing of non-canonical structures

Abstract[1]

The present study attests that a written word sequence involving a stress clash configuration and thus forcing readers to deviate from the optimal rhythmic alternation of stressed and unstressed syllables may critically hamper the processing of complex, non-canonical sentences. At points of high memory demand, the cognitive load associated with a stress clash configuration may lead to the (temporary) overwriting of memory traces necessary for the processing of syntactic long distance dependencies. I argue that this overwriting becomes possible because the clash information implicit in the written string is processed immediately while access to stored information in working memory may be tardy.

[1]This chapter will appear in the volume *Rhythm in Phonetics, Grammar, and Cognition*, edited by Ruben van de Vijver and Ralf Vogel, Berlin: Mouton de Gruyter.

2.1 Introduction

Comprehending a written sentence necessarily involves recourse to grammatical knowledge at several linguistic levels. The linearly ordered lexical items that are conveniently demarcated by blanks must be parsed into a sufficiently coherent, hierarchical syntactic representation in order to be assigned the proper meaning. At the same time, readers generate a phonological representation that they might put to use in oral reading. Unlike segmental phonology, which is relatively well represented in the orthographic code, prosody is not explicitly coded. Prosodic features such as phrasing and accentuation must be derived from, and ideally reflect, the lexical, syntactic, and focus-structural analysis of the word string.

Obviously, sentences vary in their linguistic complexity and, correspondingly, the cognitive resources needed for processing differ. More often than not, complexity at the syntactic or discourse structural level engenders complexity at the prosodic level, as indicated by the insertion of prosodic phrase boundaries or the realization of accents at positions that would remain unaccented in simple canonical structures. Moreover, independent of the effects of syntactic and discourse structure on prosodic complexity, the phonological representation of a written sentence may be complex all by itself. For segmental phonology, this is vividly demonstrated by tongue twisters and the difficulty they bring about in reading even if their syntax is relatively simple (McCutchen and Perfetti, 1982; Wilshire, 1999).

The present study is concerned with prosodic complexity resulting from deviance from the favored rhythmic alternation of stressed and unstressed syllables, focusing specifically on the effect of stress clash, i.e. the adjacency of two syllables bearing lexical stress. We report on an oral reading experiment showing that, under conditions of cognitive duress, the preference for rhythmic alternation of strong and weak syllables (cf. Schlüter, 2005; Selkirk, 1984) prohibits the proper placement of contrastive accents required for a representation in which phonological structure conforms with syntactic and focus-structural representations. The experiment suggests that preferences for local prosodic well-formedness may override more global syntactic and discourse structural constraints, and that prosodic properties of sentences may critically contribute to the cognitive load during the analysis of written text. This interpretation would be at odds with conceptions of grammar and language processing in which the phonological component is assumed to merely interpret the syntactic and focus-structural conditions.

Before reporting on the experiment in section 2, we will provide some background on the grammar of prosodic prominence and its relation to syntax and focus structure, as applied to Germanic languages (section 1.1). Section 1.2 and 1.3 introduce the experimental task, viz. unprepared oral reading, and the linguistic construction under examination, i.e. elliptic coordinations of the right-node-raising type (RNR).

2.1.1 Prosodic prominence and linguistic rhythm

Prominence patterns in oral language emerge from the interplay of various forces. The lexicon specifies which syllable of a given word is assigned main stress and which syllables receive secondary or no stress. There is no clear phonetic correlate of stress – its phonetic realization crucially depends on the prosodic context. In general however, stressed syllables show longer duration, higher intensity and more pitch modulation than unstressed ones (Beckman and Pierrehumbert, 1986a). On the supra-lexical level, syntactic and focus-structural conditions determine which words receive phrase or sentence accent. Accent is realized by a clear deflection in pitch on or near the stressed syllable of the designated word. Phrase accent is thus more prominent than word stress. If a sentence like (1-a) is uttered out of the blue, the words 'boy' and 'biscuits' typically receive accent (lexical stress is marked by underlining, accent is marked by small caps). 'Biscuits' is assigned the strongest prominence in the sentence (the nuclear accent), which is realized as a pitch accent on or near the syllable carrying main lexical stress (the first syllable in the case of 'biscuits'). Nuclear accent on the verb (1-b) is possible only under certain discourse structural conditions, i.e. if the accented verb is contrasted with some other predicate in the context, as would be the case if (1-b) was uttered as a clarification to (1-a).

(1) a. [$_S$ [$_{NP}$ The <u>lit</u>tle <u>BOY</u>] [$_{VP}$ <u>likes</u> [$_{NP}$ <u>BIS</u>cuits.]]]
    b. [$_S$ [$_{NP}$ The <u>lit</u>tle <u>boy</u>] [$_{VP}$ a<u>DORES</u> [$_{NP}$ <u>bis</u>cuits.]]]

The literature on sentence phonology attributes such a pattern to the workings of the syntax-phonology interface (Selkirk, 1995; Truckenbrodt, 2006, 2007). Here, we follow an Optimality Theoretic (OT) account in which accent assignment is captured by the interaction of violable constraints regulating the mapping between i) focus structure and phonology, and ii) syntax and phonology. The latter require the assignment of accents to lexically headed XPs (StressXP) and furthermore demand that the rightmost accent be strongest (Rightmost). Within the verb phrase, accent on the object argument fulfills StressXP for both the object NP and the VP as a whole since the NP is a proper constituent of the VP. Accent on the verb without accent on the object, however, would violate StressXP due to the lack of accent on the NP constituent (cf. Truckenbrodt (2007) for details of this analysis).

(2) a. StressXP: Each lexically headed XP contains an accent.

b. RIGHTMOST: The rightmost accent within a prosodic or intonational phrase is strongest.[2]

If (1-a) provides the context for a clarification in the form of a statement like (1-b), accent on the verb is required in order to reflect the contrast of the verbs and the givenness of the remaining sentence. This focus-structural condition can be captured by the constraint STRESS-FOCUS (Féry and Samek-Lodovici, 2006; Samek-Lodovici, 2005; Selkirk, 1995; Zubizarreta, 1998), which, in a nutshell, demands accent on contrasted or focused material. As evidenced by (1-b), STRESS-FOCUS may override syntax-driven accent assignment.

(3) STRESS-FOCUS: A focused phrase has the highest prosodic prominence.

Matters are complicated by a general preference for rhythmic alternation of strong (i.e. stressed) and weak (i.e. unstressed) syllables (Alber, 2005; Hayes, 1995; Kager, 1989; Liberman and Prince, 1977; Schlüter, 2005; Selkirk, 1984). This preference (formulated in OT terms as the constraint *CLASH) operates against stress clashes (i.e. sequences of adjacent syllables bearing lexical stress), which are avoided whenever more rhythmic alternatives are available.

*CLASH may require readjustments of lexically specified stress patterns to allow for rhythmic alternation of strong and weak syllables, as exemplified in the phrase *ideal* + *partner* → *ideal* PART*ner*.

However, such readjustments are clearly restricted. In the case of (1-a) the stress clash between the verb 'like' and the following object 'biscuits' cannot be resolved by shifting the accent onto the last syllable of the object. Instead, the stress clash is tolerated in (1-a). As a rule, metrical readjustments triggered by *CLASH do not concern words that bear main phrase or sentence accent but may only affect unaccented (or more weakly accented) neighboring words (Grabe and Warren, 1995; Hayes, 1995) (but see Berg (2008) for emphatic stress shift). In other words, accents that are assigned due to STRESSXP may not be altered by the desire for rhythmic alternation.

Independent of accentuation, stress shift is also blocked if a prosodic phrase boundary intervenes between two adjacent stressed syllables. This is typically the case between the subject and the verb phrases in sentences like (4) (Selkirk, 1995). Here, the gerund phrase in subject position is separated from the verb phrase by a prosodic phrase boundary. Consequently, stress on the following monosyllabic verb does not trigger stress shift on 'TV' .

(4) (Watching CABLE tv) (harms children's health).

[2]This constraint is equivalent to HEAD PHRASE (HP/HI) (Féry and Samek-Lodovici, 2006), which demands that heads of prosodic or intonational phrases appear close to their right edge.

In summary, accent assignment driven by focus may override syntax-driven accent assignment (as exemplified in (1)), which in turn restricts the application of rhythmic readjustments. We will therefore subscribe to the hierarchy of constraints affecting the location and patterning of prosodic prominences formulated in (5):

(5) Stress-Focus ≫ StressXP ≫ *Clash

The hierarchy in (5) corresponds with the size of the domains over which the constraints exert their influence. Stress-Focus evaluates the discourse context while StressXP acts on a considerably smaller level, i.e. the level of the syntactic XP. *Clash only takes local information into account, operating on the level of the syllable sequence. Accordingly, the grammar of prosodic prominence is shaped by the interaction of weak local with strong global constraints. This state of affairs has interesting psycholinguistic implications.

Language processing proceeds in an incremental way. That is, the language processing mechanism builds linguistic representations based on the piecemeal access to the input as soon as it becomes available (cf. Pickering and van Gompel (2006) and references therein). In the case of a contextless sentence, the parser initially has access only to local (lexical) information – the sequence of syllables or words that have to be integrated into more global representations such as syntactic phrases. The propositions expressed in phrases or sentences then build the base for the discourse setting. Hence, information access in (contextless) language comprehension proceeds from local information to global information.

Following this rationale, a constraint like *Clash, its weak role in the grammar notwithstanding, may have a more immediate influence on incremental structure building compared to constraints like StressXP and Stress-Focus that operate on larger domains. Ergo, there is a certain tension between the hierarchy of constraints in grammar on the one hand and the order of access to their respective domains in language comprehension on the other. In view of these considerations, we hypothesize that *Clash, despite its low rank in grammar, may have a crucial role in language processing in affecting early stages of structure building.

Note that the constraints in (5) are geared towards the assignment of prosodic structure, i.e. the pattern of prosodic prominence. They are therefore irrelevant for auditory sentence comprehension, in which prosody is already an integral part of the input. In auditory language comprehension, prosody has to be interpreted, not assigned to the input (cf. Cutler et al. (1997) for the role of prosody in auditory sentence processing). Processing written language, however, does involve prosodification of the input string. This is especially obvious in the case of oral reading.

2.1.2 Oral reading

Fluent oral reading involves both sentence comprehension and sentence production. As for comprehension, readers have to parse the word string into a sufficiently coherent syntactic representation in line with the discourse context. At the same time, readers produce speech using prosody that conforms to the syntactic analysis and the focus-structural setting of the text (Koriat et al., 2002; Kreiner and Koriat, 2005; Kondo and Mazuka, 1996; Wheeldon, 2000).

The simultaneity of language comprehension (syntactic parsing) and production (prosodification) in oral reading suggests that these processes might interact. The role of prosody is especially intricate in this task: there is no overt correlate of prosody in the graphemic string.[3] As neither syllables nor lexical stress or accent are marked orthographically, readers have to deduce these phonological features on the basis of i) the lexical and syntactic information derived from the word string, and ii) the focus-structural representation extrapolated from the syntactic and semantic analysis of the text. Unprepared readers have been shown to produce prosody in accordance with these representations (Koriat et al., 2002; Kreiner and Koriat, 2005). The conformity of reading prosody with the syntactic and focus-structural facts indicates the dependence of prosody on the syntactic and focus-structural analysis. Also, in research on reading development, the prosodic appropriateness of read text is used as a diagnostic for reading comprehension and reading skill in general (Schwanenflugel et al., 2004). This would suggest that reading prosody is only constructed on the basis of considerable syntactic and semantic pre-processing.

On the other hand, there is evidence that prosody derived from the written string is used by readers to make syntactic parsing decisions. That is, prosody constructed during reading is recycled and used in a way similar to prosody in auditory sentence comprehension (Fodor, 2002) where it might disambiguate otherwise ambiguous word strings. Various reading studies have demonstrated the impact of implicit prosodic factors on sentence comprehension in silent reading (Augurzky, 2006; Bader, 1998; Fodor, 1998, 2002; Hwang and Schafer, 2009; Stolterfoht et al., 2007). These studies are concerned with the role of phrase accent and prosodic phrasing in resolving syntactic ambiguities. Results by (Kentner, 2012, cf. chapter 3 of this dissertation) suggest a role of stress-based linguistic rhythm in sentence comprehension, too. In the face of a syntactically ambiguous word string that involves a stress clash in one reading but not in the other, the parser favors the

[3]Commas are not a reliable, undisputed cue to prosodic phrasing either. Chafe (1988) argues convincingly that the commas in the phrase *'red, white, and blue'* do not correlate with prosodic breaks, while the commas in *'Abernathy came, Chippendale saw, and Higginbottom conquered'* do. Punctuation rules do not refer to phonological weight or phrase length, but prosodic phrasing does.

analysis of the rhythmically nonoffending reading.

Therefore, instead of a unidirectional dependence of prosody on the syntactic and focus-structural analysis, there is good reason to assume an interrelationship of prosody with these representations in reading.

The fact that reading prosody gives an insight into the syntactic and focus-structural analysis makes oral reading a perfect test environment for examining the interaction of phonological, syntactic, and discourse structural processes in language processing. For this, we study RNR-type elliptic coordination structures, which feature an interesting interplay of syntax, focus structure and prosody.

2.1.3 The right-node-raising construction

Focus structure and prosody of RNR

The so-called right-node-raising construction[4] (henceforth: RNR) is a type of elliptic coordination structure in which an element that overtly appears at the right periphery of the second conjunct is understood as part of both the first and second conjuncts. This element, i.e. the target of ellipsis, is represented by crossed-out letters in (6).

(6) Peter kauft ~~Kekse~~ und Hans isst Kekse.
Peter buys ~~biscuits~~ and Hans eats biscuits.
Peter is buying and Hans is eating biscuits.

Hartmann (2000) and Féry and Hartmann (2005) formulate the conditions that must be met in order for RNR to be applicable. First, the element preceding the target of ellipsis has to be stressable. Hence, the modal verb in (7) must not appear in its reduced form, otherwise RNR is impossible.

(7) I think that {I would / *I'd} and I know that {he will / *he'll} buy the pictures.

Secondly, the conjuncts must exhibit a parallel syntactic and focus structure and the pre-elliptic parts of the two conjuncts must allow for a contrastive interpretation.

[4]Following Hartmann (2000), Phillips (1996), and Wilder (1997), we consider RNR to involve deletion rather than raising of the target of ellipsis. Various facts speak against a movement analysis of RNR, chief among them the observation that nonconstituents may be the target of ellipsis in this construction (cf. (i)).

(i) Peter verspricht seinem [ ] und Maria verspricht ihrem [Kind ein Geschenk].
Peter promises his [ ] and Maria promises her [child a present].
Peter promises a present to his child and Maria promises a present to her child.

Correspondingly, RNR sentences display a complex focus structure with contrastive focus embedded in a broad presentational focus (Selkirk, 2002; Féry and Samek-Lodovici, 2006). Since they represent material that is new to the discourse, the two conjuncts including the target of ellipsis are instances of broad presentational focus. Moreover, the pre-elliptical elements in both conjuncts are semantically contrasted and correspondingly bear contrastive focus. We follow Selkirk (2002) in marking broad presentational focus by *foc* and contrastive focus by Foc. In (8) the focus structure with embedded contrastive focus is represented for (6).

(8) [Peter [kauft]$_{FOC}$ ~~Kekse~~]$_{foc}$ und [Hans [isst]$_{FOC}$ Kekse]$_{foc}$

The focus structure of (8) thus differs from the nonelliptical counterpart with the same basic constituent order in (9). This structure may represent a contrast between the two conjuncts. The two verbs, however, cannot contrast in (9) as they differ with respect to transitivity.

(9) [[Peter lacht]$_{FOC}$und [Hans isst Kekse]$_{FOC}$]$_{foc}$
Peter is laughing and Hans is eating biscuits.

The focus structure determines the prosodic realization of the RNR sentence, which differs from the prosody of comparable nonelliptic sentences. As discussed above, contrastively focused elements are assigned prosodic prominence due to Stress-Focus. Accordingly, in (8) the contrastively focused verbs are accented, while in (9), StressXP determines the position of accent in the second conjunct, i.e. nuclear stress falls on the object. While production studies on RNR prosody (Féry and Hartmann, 2005; Kentner, 2007; Kentner et al., 2008; Selkirk, 2002) may differ on details of the prosodic phrasing of RNR, they agree that the pre-elliptic element is the location of the nuclear accent.

The following Tableau reflects these basic facts as derived from the grammatical interface of syntax and information structure with prosody.

/[H. isst Kekse]$_{foc}$/		Stress-Focus	StressXP
a.	☞ H. isst KEKSE		
b.	H. ISST Kekse		*!
/[H. [isst]$_{FOC}$ Kekse]$_{foc}$/		Stress-Focus	StressXP
a'.	H. isst KEKSE	*!	
b'.	☞ H. ISST Kekse		*

Table 2.1: Prosody of 2^{nd} conjunct in nonelliptic (9) and RNR sentences (8).

Processing RNR

There is currently only little psycholinguistic evidence concerning the processing of RNR sentences. Processing RNR sentences can be considered a relatively complex task because of the required nonlocal interpretation of the elliptic target. As for auditory sentence comprehension, processing difficulty due to the ellipsis may be alleviated by the characteristic RNR prosody. The contrastive accent on the verb, deviating from typical nuclear accent in nonelliptic contexts, signals the contrast and, correspondingly, the ellipsis. However, as Kentner et al. (2008) find, processing ease depends not only on the contrastive accent but also on the strength of the prosodic break at the conjunction. Listeners were shown to benefit from a clear prosodic break before the conjunction, a type of phrasing that speakers seem to avoid. Since prosody is not provided in the written modality, processing the RNR-type ellipsis in reading may turn out to be particularly difficult. To our knowledge, there is no study as yet examining RNR in written language processing.

We may however conjecture the following concerning the processing of RNR structures in reading: In the case of a RNR sentence like (8), the ellipsis is implicitly marked by the fact that the argument structure of the transitive verb 'kauft' in the first conjunct is not satisfied locally. Assuming that the reader observes the imperative of parallelism in coordinations, the parser could, already at the conjunction, predict a verb phrase or a complete sentence to form the second conjunct. Moreover, the only way to construct a fully grammatical coordination structure is to project a transitive VP, the object of which simultaneously functions as the object of the first conjunct. Hence, already at the conjunction, readers might recognize the ellipsis and predict the contrast that is characteristic of RNR. Note, however, that forming such a long-distance dependency imposes high processing costs due to the strain on memory capacity (e.g. Hawkins, 1994). If the parser can bear these costs, readers should be able to assign contrastive accent to the verb in the second conjunct. However, the high strain on the processing mechanism might make the analysis relatively vulnerable.

In fact, various studies have demonstrated the difficulty associated with processing noncanonical structures that induce high memory costs (cf. the literature on nested structures like center embedding). In such a situation, the parser might resort to some sort of shallow processing (Ferreira, 2003; Sanford and Sturt, 2002). Shallow processing implies ignorance towards structural information that would be required in order to form a fully specified grammatical representation. Instead the parser will make do with less than perfect comprehension. Gibson and Thomas (1999) put forward the 'high memory cost pruning hypothesis' stating that syntactic predictions incurring the most memory load are forgotten if processing costs exceed a certain threshold, i.e. at points of high memory complexity during the parse.

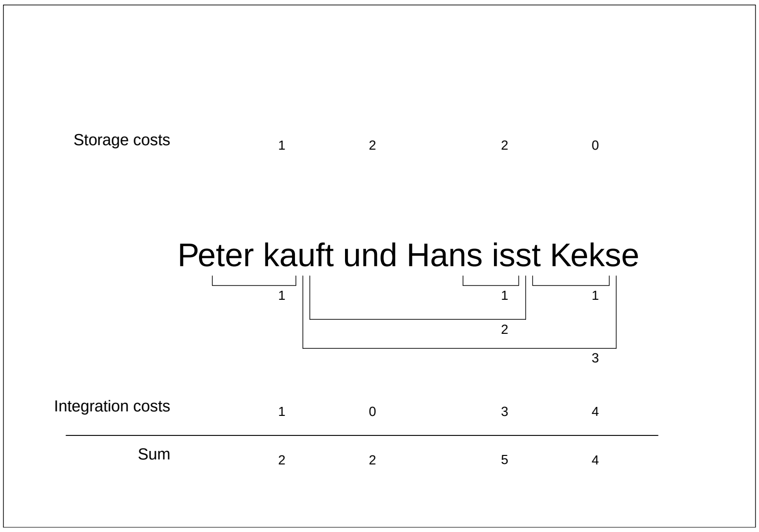


Figure 2.1: DLT predictions for (8) (Translation: *Peter is buying and Hans is eating biscuits.*). The highest processing costs are predicted for *isst* in the second conjunct.

Dependency Locality Theory (DLT, Gibson, 2000) is used here to determine the point associated with the highest memory complexity in parsing RNR structures.[5] There are two components to the DLT: a storage cost component and a prediction cost component. Both components contribute additively to the cognitive costs at any given point in a sentence. According to the prediction component, each syntactic head that needs to be predicted in order to complete the current input as a grammatical sentence incurs memory costs. According to the storage cost component, memory costs increase with every new discourse referent intervening between a word and the dependent to be integrated with it.

Concerning RNR sentences like (8) (cf. Figure 1), the elliptic first conjunct incurs relatively high prediction costs, in that a second, transitive VP including an object NP needs to be projected (two syntactic heads) once the conjunction is reached. The point of greatest difficulty, however, is at the verb in the second conjunct (*isst*). At this position, the second verb needs to be integrated with its preceding subject *Hans*. Furthermore, the contrast to the verb in the first conjunct needs to be computed across two new discourse referents (*Hans* and *isst*). Also, two NPs are predicted at *isst*, i.e. the coincident object of the first and the second conjunct.

With these considerations in mind, we turn to the role of rhythmic prefer-

[5] Other distance-based accounts of syntactic and parsing complexity are formulated in Hawkins (1994) and Joshi (1990).

ences and their interaction with other constraints regulating prosodic prominence in written sentence processing. Given the rather weak standing of *Clash in the grammar of prosodification, one might ask whether this constraint is merely a stylistic force or whether it has a functional role beyond the amelioration of prosodic structure. As stated above, *Clash may be considered important as it evaluates relevant features of the written input more immediately than the concurring constraints StressXP and Stress-Focus, simply because the relevant domain, i.e. syllables, are among the earliest information available to the processing mechanism (Ashby and Martin, 2008; Ashby and Rayner, 2004). As discussed above, language processing proceeds incrementally, so local information may be more readily available to the processor than higher level information, which necessitates an overview of the more global context. Also, global information that is based on long-distance dependencies (i.e. distant memory traces) is likely to be forgotten by a parser that has to cope with high processing demands.

Before we examine the role of *Clash experimentally, we make the following assumptions: In speech production, selecting the points of prosodic prominence (accents) is more difficult when elements that may function as the target of prosodic prominence (stressed syllables) cluster together. We therefore conjecture that the formation of a mental representation involving a violation of *Clash is cognitively costly and may thus hamper processing (see also the neurophysiological account of the cognitive complexity of stress clash configurations by Schlüter (2003)).

Likewise, producing accents at positions that would remain unaccented in canonical sentences is considered a costly operation (cf. Reinhart, 2006): That is, a violation of StressXP, which is mandatory in the type of RNR under study here, causes cognitive duress. Corresondingly, the cognitive costs associated with reading/producing a RNR sentence are higher than those associated with producing nonelliptic coordinations. Arguably, stress clash configurations increase the cognitive load associated with RNR. According to Gibson's memory complexity metric, the analysis is most vulnerable at the second verb. We study the influence of *Clash at this critical position.

2.2 Experiment

In order to scrutinize the effects of StressXP and Stress-Focus and their interaction with *Clash in reading, we use RNR and comparable nonelliptic coordinations as a test environment. Juxtaposing RNR and nonelliptic coordinations makes the workings of Stress-Focus and StressXP transparent (cf. the Tableau above). Varying the rhythmic environment allows us to gauge the influence of *Clash in reading.

Given the close compliance of accentuation and contrast that is formulated in the grammatical constraint STRESS-FOCUS, reading prosody is a good way to test whether readers indeed form the contrast and correspondingly parse the ellipsis. We therefore use the realization of accent as the dependent variable for determining whether readers process a valid RNR.

2.2.1 Method

Design and material

The objects of investigation are elliptic (RNR) and nonelliptic coordinations with the same basic constituent order. Aside from prosody, the verb in the first conjunct distinguishes elliptic and nonelliptic versions. The latter are characterized by an obligatorily intransitive verb while the former are identified by a transitive verb in the first conjunct. Prosodically, the versions differ with respect to the location of nuclear accent in the VP of the second conjunct. Nonelliptic versions feature nuclear accent on the object while elliptic versions display nuclear accent on the verb. The rhythmic environment is varied on the words preceding and following the critical verb in the second conjunct inducing either rhythmically alternating sequences or stress clashes to either side. The subject preceding the verb in the second conjunct is trisyllabic with either initial or final stress. The object is minimally disyllabic with lexical stress either on the initial or the second syllable.

The experiment has thus a 2x2x2 factorial design with the factors 'ellipsis' (RNR vs. nonelliptic), 'rhythmic environment to the left (ClashL)' (trisyllabic subject with initial or final stress) and 'rhythmic environment to the right (ClashR)' (object with initial or noninitial stress). For this, 28 sets of sentences in eight conditions were constructed. Table 2 shows an example set (the full list of experimental sentences is listed in the appendix).

Participants

24 female first year undergraduate students from the University of Potsdam took part in the experiment. All are native speakers of German and naïve as to the purpose of the experiment. They either received course credit or were paid 5 Euros for their participation.

Experimental procedure

The experimental sentences were divided into four lists using a Latin square design such that conditions were maximally counterbalanced across lists and each participant would see at most two sentences from each of the 28 sets. The items were fed into a DMDX presentation (Forster and Forster, 2003) together with 65 unrelated

	Ellipsis	ClashL	ClashR	example
a.	-1	1	1	Karl lacht und der Dirigent isst KUCHEN.
b.	-1	-1	1	Karl lacht und der Musiker isst KUCHEN.
c.	-1	1	-1	Karl lacht und der Dirigent isst GEBÄCK.
d.	-1	-1	-1	Karl lacht und der Musiker isst GEBÄCK.
e.	1	1	1	Karl holt und der Dirigent ISST Kuchen.
f.	1	-1	1	Karl holt und der Musiker ISST Kuchen.
g.	1	1	-1	Karl holt und der Dirigent ISST Gebäck.
h.	1	-1	-1	Karl holt und der Musiker ISST Gebäck.

Translation: *Karl is* {a-d: *laughing*; e-h: *bringing*} *and the* {a,c,e,g: *conductor*; b,d,f,h: *musician*} *is eating* {a,b,e,f: *cake*; c,d,g,h: *pastries*}.

Table 2.2: Factors with coding and sentence materials (example). Stressed syllables are underlined, expected accent is marked by small caps.

fillers and pseudo-randomized for each subject using the Mix randomization tool (van Casteren and Davis, 2006) such that sentences of the same condition had a minimum distance of eight and sentences of the same experiment had a minimum distance of three items. The experiment was set up as an unprepared reading design in which participants read each sentence aloud without advance preparation as soon as it appeared on screen.

The experiment took place in an acoustically shielded room with an AT4033a studio microphone. Each participant was seated in front of a 15" computer screen with the microphone placed approximately 30 cm from the participant's mouth. A keyboard was placed on a table in front of the subject. Recordings were made on a computer using the RecordVocal function of DMDX and a C-Media Wave sound card at a sampling rate of 44.1 kHz with 16 bit resolution. The DMDX presentation was programmed for each item as follows: First, only the first one or two words (the sentence initial subject noun phrase or proper name) were presented on the screen. Participants were told to familiarize themselves briefly with these words. They were instructed to then press the space bar, inducing the presentation of the entire sentence. Participants were asked to start reading out the sentence as soon as it appeared on screen and to do so as fluently as possible. The spacebar press automatically initiated the recording. After a fixed recording time of five seconds, the procedure was repeated for the next item. For each sentence, only one realization per subject was recorded. No corrections were recorded in the case of hesitations or slips of the tongue.

Data analysis

All in all, (24*28=)672 experimental sentences were recorded. The sentences were independently judged by two students who were blind to the conditions and the purpose of the experiment. Their task was i) to note slips of the tongue and disfluencies, and ii) to determine for each sentence whether nuclear accent was realized on the verb or on the object. The judges were paid for their work. 85 sentences (13%) were marked by at least one of the judges as nonfluent or as containing slips of the tongue. Concerning the position of nuclear accent, the judges agreed on 510 of the 587 fluent sentences (87%). The sound files of the 587 flawless sentences were hand-annotated by a phonetically trained student who was blind to the purpose of the experiment and to the judgments of her fellow students. The second conjunct was segmented into words and syllables and labeled accordingly. In the following section, we report the results from the set of consistently judged sentences. Separate analyses for the two judges yield comparable results.

2.2.2 Results

Flawed sentences

The number of flawed sentences is relatively high (n=85, 13%), which can be partly explained by the task (unprepared reading) and the length and complexity of the sentences. In order to check whether the distribution of flawed sentences is systematically related to the controlled factors of the experiment, we fitted a generalized linear model with binomial link function (Bates and Sarkar, 2007; Gelman and Hill, 2007; Quené and van den Bergh, 2004). The fixed factors of this model are i) 'ellipsis' (elliptic vs. nonelliptic), ii) 'the rhythmic environment to the left' (initial vs. final stress on the subject of the second conjunct), and iii) 'the rhythmic environment to the right' (initial vs. noninitial stress on the object of the critical verb). Flawed versus fluent realization was used as the binomial dependent variable; variance due to individual participants and items was taken into account by including these factors as grouping variables. In order to avoid correlations of the fixed factors in the statistical models, orthogonal or contrast coding was applied (factor 'ellipsis': elliptic=1, nonelliptic=-1; factor 'ClashL': clash=1, no clash=-1; factor 'ClashR': clash=1, no clash=-1). No significant effect was found for either of the fixed factors, nor for the interaction (all z-values are distinctly < 2) suggesting that the controlled variables do not systematically influence the distribution of flawed sentences.

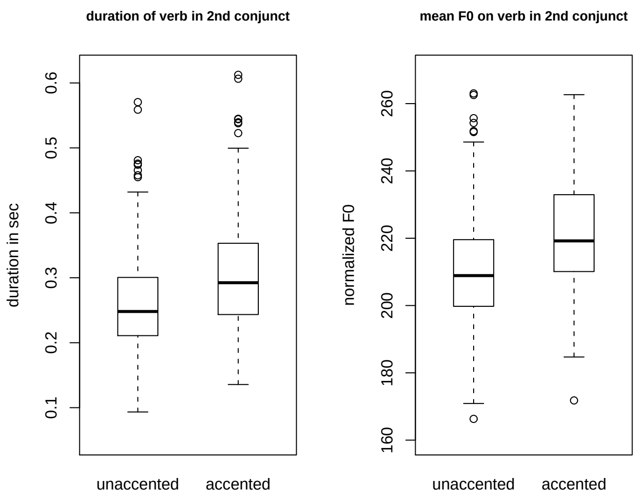


Figure 2.2: Box plots showing median and interquartile ranges for duration (left panel) and normalized mean F0 (right panel) of the critical verbs in the second conjunct, broken down by perceived accentuation.

Phonetic analysis of judgment data

A phonetic validation of the judgments on accentuation is required, as listeners may perceive prominence patterns on syllable sequences in context even in the absence of definite acoustic cues for them (e.g. Dilley and McAuley, 2008). Syllable durations and mean pitch on the critical verb in the second conjunct were compared for realizations with perceived nuclear accent on the object and nuclear accent on the verb. The F0 values for the verb were normalized prior to analysis. The normalizing factor used is the mean F0 across speakers on the verb divided by the utterance wide mean F0 of each individual sentence.

The results reveal longer durations and a higher mean F0 for those verbs that were perceived as bearing nuclear accent. Verbs in this condition were on average 50 ms longer and around 11 Hz higher in mean F0 compared to unaccented versions. These differences are comparable to the ones reported in the literature on the acoustics of accented versus unaccented syllables (Eady et al., 1986; Baumann et al., 2006; Féry and Kügler, 2008).

Linear mixed effects models confirm the phonetic difference between verbs with versus without perceived nuclear accent. The dependent variables 'mean F0' (in Hz) and 'duration' (in ms) were logarithmized for inferential statistics to adjust for the skew in the data. Using participant and item as random effects, the model

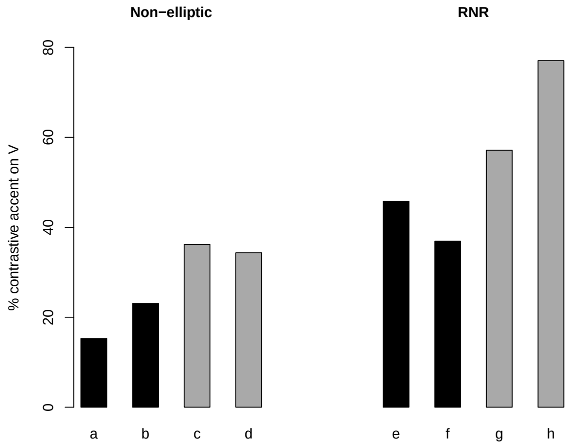


Figure 2.3: Bar plot showing percentages of verbs bearing contrastive accent in nonelliptic versus elliptic (RNR) sentences by experimental condition. Black bars represent conditions with stress clash configuration between the verb and following object (ClashR=1).

evaluating the duration of the verb against the perceived accentuation reveals a significant effect (Coefficient estimate: 0.129, SE: 0.017, t-value[6]=7.487). A comparable model with logarithmized mean F0 as dependent variable yields a complementary effect (Coeff. estimate: 0.0437, SE: 0.0093, t-value=4.7). The phonetic analyses thus confirm the reliability of the judgment data on the perception of nuclear accent.

Placement of nuclear accent

In total, the verb in the second conjunct was perceived as bearing nuclear accent in 40% of the cases. As expected, readers produced significantly more accents on the verb in RNR conditions (on average 54% of the accents were on the verb in RNR conditions). However, in roughly 27% of the cases, readers produced a nuclear accent on the verb in nonelliptic sentences. Conversely, in 46% of the cases readers failed to produce a nuclear accent on the verb where it was required. Crucially, aside from the syntactic construction, the distribution of accents is also determined by the rhythmic environment. Fewer nuclear accents were realized on the verb when they would clash with adjacent stressed syllables. Readers produce particularly few nuclear accents on the verb if the following object bears

[6] t-values greater than |2| mean statistical significance at the level of α=0.05.

initial stress (roughly 13% fewer accents as compared to rhythmically unoffending environments). The bar plot shows the results concerning the judgments on the fluent sentences. The bars represent percentages of perceived nuclear accent on the verb in the second conjunct broken down by elliptic (right panel) versus nonelliptic (left panel) conditions. A generalized linear model with participants and items as random effects evaluating perceived nuclear accent position against the crossed fixed factors confirms the significant effects for the factor ellipsis and the rhythmic environment to the right. The effect for the left environment and the interactions remain nonsignificant.

Coefficient	Estimate	Std. Error	z-value	p-value
Ellipsis	0.8707	0.1208	7.209	<0.001*
ClashRight	-0.6789	0.1194	-5.686	<0.001*
ClashLeft	-0.1583	0.1168	-1.356	0.175
Ellips:ClashR	-0.1105	0.1160	-0.952	0.341
Ellips:ClashL	-0.1271	0.1182	-1.076	0.282
ClashR:ClashL	0.2298	0.1982	1.159	0.246
Ell:ClashR:ClashL	0.1539	0.1550	0.993	0.321

Table 2.3: Results of generalized linear model evaluating perceived nuclear accent against crossed fixed factors.

Summarizing the results, we find that i) readers realized contrastive nuclear accent on the verb more often in sentences that were devised as RNR, and ii) that fewer nuclear accents were realized on the critical verb if the following object had lexical stress on the initial syllable.

2.3 Discussion and conclusion

The results confirm that the manipulation concerning the transitivity of the verb in the first conjunct and, correspondingly, the ellipsis did work. Readers were clearly more prone to produce a contrastive nuclear accent on the second verb when the argument structure of the first conjunct was incomplete. This suggests that readers realize the presence of an ellipsis.

However, not all readers consistently produced contrastive accents in such situations, verifying that the sentence type (RNR) and the task (unprepared reading) are cognitively demanding. A key reason for the failure to contrastively accent the second verb in elliptic sentences – besides the difficulty of parsing a long-distance dependency in unprepared reading – lies in the rhythmic environment, as confirmed by the significant influence of the stress location on the object (ClashR) on accent placement. Initial stress on the object together with the stressed verb induces a stress clash configuration. But why should the clash be responsible for the failure to realize contrastive accent on the verb? The possible answer lies in the cognitive costs associated with stress clashes and the immediacy of their evaluation by the parser. As hypothesized at the outset, producing noncanonical nuclear accents in a stress clash environment is cognitively costly: potential targets of the accent (stressed syllables) cluster together and hence complicate the selection of the grammatically prescribed accent position. The clash thus adds to the cognitive demands associated with processing the long-distance dependency. The violation of *Clash might be tipping the scales and forcing the reader to drop the RNR analysis in favor of a nonelliptic analysis with nuclear accent on the object. That is, the processing costs associated with the stress clash make readers ignore (or forget about) the incomplete argument structure in the first conjunct, which otherwise makes them predict and, correspondingly, realize the contrast. This interpretation of the results is in line with the 'high memory cost pruning' hypothesis (Gibson and Thomas, 1999). In forgetting about the structural dependency between the elliptic first VP and the VP in the second conjunct, the parser gets rid of the most costly memory traces at points of excessive cognitive complexity. As a consequence, the ellipsis is ignored and the second conjunct is processed as it would be in nonelliptic coordinations, with nuclear accent on the object. The results might also be interpreted in light of recent work on grammatical illusions, i.e. sentences which are acceptable (and appear to be grammatical) but turn out to be ungrammatical on closer inspection.[7] As Haider (2011) notes, grammatical illusions are characterized by local wellformedness but are globally deviant. Apparently, this characterization also holds for the RNR sentences that were realized without contrastive accent on the verb (10). The prosodic rendition in (10) fails to mark the contrast between the verbs that is a prerequisite for the ellipsis.

(10) #Hans kauft und Peter isst Kekse.
Hans is buying and Peter is eating biscuits.

However, the second conjunct is perfectly wellformed as long as the requirements of the first conjunct are disregarded. The transitive VP *isst Kekse* bears nuclear accent on the object, just as would be expected in canonical transitive VPs. Conversely, the required contrastive accentuation on the verb is less wellformed locally (cf. the violation Stress-XP and, if applicable, *Clash) and licit only under global pressure (satisfaction of Stress-Focus).

In this context, it is important to remember that readers could already predict the ellipsis and hence the requirements of Stress-Focus at the conjunction on

[7]A famous example of a grammatical illusion is the utterance *More people visited Vienna than I did.* Various examples can be found in Haider (2011).

the basis of the unsatisfied argument structure of the first verb. Lookahead should be critical to sentence comprehension, since readers need to use it in order to construct a syntactically wellformed and coherent analysis of the written string online. Crucially, the extent and potency of the lookahead is significantly constrained by the local rhythmic environment.

Interestingly, the rhythmic environment to the left of the critical verb does not seem to affect accent realization. This lack of an effect is explicable with recourse to the phonological phrasing that comes with the syntactic structure of the second conjunct. The constituent to the left of the critical verb is the subject of the second conjunct. The subject projects its own XP and, according to standard assumptions on the syntax-prosody interface, might therefore be separated from the verb phrase by a phonological phrase boundary. As evidenced by (4) above, the adjacency of stressed syllables is generally tolerated by *Clash if a prosodic boundary intervenes. This suggests that the syntax-driven insertion of prosodic phrase boundaries took place before the stress clash could hamper processing. It seems that the processing costs normally associated with stress clash are not incurred in the context of prosodic phrase boundaries.

The rhythmic effect to the right of the critical verb indicates that *Clash, which is traditionally understood as a production-oriented constraint, may interfere with comprehension when violated. The cognitive complexity of stress clashes is explained with recourse to production: producing accents in stress clash environments is complex because selecting the appropriate syllable for accentuation is difficult when potential targets cluster together. The results of the experiment, however, suggest that stress clash configurations in written text influence comprehension. The cognitively costly stress clash representation critically aggravates the already complex focus-structural and syntactic analysis of elliptic sentences.

This state of affairs is hardly compatible with the conception of a uni-directional relation between syntactic and prosodic processes in written sentence comprehension. Rather, it calls for a more interactive view of these processes. The experimental results confirm the suggestion that the simultaneity of comprehension and production in oral reading indeed implies an interrelationship. This is also in line with findings from silent reading that have shown the significance of rhythmic effects for sentence comprehension (Kentner, 2012, cf. chapter 3 of this dissertation).

Aside from these principled considerations concerning the architecture of the human language processing mechanism and its access to different domains of grammar, the current experiment raises methodological issues: The complexity of, and hence the predicted parsing difficulty associated with experimental sentences is standardly manipulated at the level of syntactic structure in psycholinguistic research. This study illustrates that prosodic properties of sentences may critically contribute to the cognitive complexity and thus should be taken into account in experiment design.

In summary, the most important conclusion to be drawn from this experiment is that forced deviance from rhythmic alternation may affect readers' performance on noncanonical sentences at points of extreme cognitive complexity. In such situations, the violation of merely weak prosodic constraints may lead to fatal pruning of memory traces that are necessary for the full syntactic and focus-structural representation and, hence, to misinterpretation.

Chapter 3

Linguistic rhythm guides parsing decisions in written sentence comprehension

Abstract[1]

Various recent studies attest that reading involves generating an implicit prosodic representation of the written text which may systematically affect the resolution of syntactic ambiguities in sentence comprehension. Research up to now suggests that implicit prosody itself depends on a partial syntactic analysis of the text, raising the question of how implicit prosody contributes to the parsing process or whether it merely interprets the syntactic analysis. The present reading experiments examine the influence of stress-based linguistic rhythm on the resolution of local lexical-syntactic ambiguities in German. Both speech production data from unprepared oral reading as well as eye-tracking results from silent reading demonstrate that readers favor syntactic analyses that allow for a prosodic representation in which stressed and unstressed syllables alternate rhythmically. The findings contribute evidence confirming immediate and guiding effects of linguistic rhythm on the earliest stages of syntactic parsing in reading.

[1]This chapter is published in *Cognition 123 (1)*, 2012.

3.1 Introduction

When reading silently, many readers experience an 'inner voice' that conveys from the graphemic string an intrinsic auditory version of the text. This mental representation has been described as being essentially speech-like, not only entailing segmental phonological information, but also prosody and even paralinguistic characteristics such as voice quality and speech tempo (Chafe, 1988). While there is little disagreement about the existence of the 'inner voice' phenomenon, it is debated whether and how the prosodic characteristics of the implicit phonological representation affect sentence comprehension.

In the present study, we will focus on one aspect of this mental representation, namely on the linguistic rhythm that emerges from the implicit stress patterns of the word sequence. The findings of two reading experiments presented here confirm that readers mentally construct patterns of implicit lexical prominences, which evolve from the concatenation of individual words. In the face of a temporal syntactic ambiguity, readers preferably generate a parse that conforms to rhythmic well-formedness principles. Specifically, the findings indicate that the initial stages of the syntactic parsing process are sensitive to the local prosodic environment even in the written modality, where no explicit prosodic cues exist.[2]

In the following section, we will briefly review findings on the relevance of linguistic rhythm in auditory language processing and discuss existing research on the role of implicit prosody for written sentence comprehension. Together, these findings motivate the two reading experiments, which are designed to shed light on the interplay of linguistic rhythm and syntactic parsing in reading.

3.1.1 Stress and linguistic rhythm in auditory language

Stress is hierarchical in the sense that for each content word there is a single syllable that carries the main stress (Hayes, 1995). Other syllables within the word either bear secondary stress or remain unstressed. Although there is no clear phonetic correlate of stress, stressed syllables are usually lengthened and may be realized with a higher pitch compared to unstressed syllables (Beckman and Pierrehumbert, 1986b; Hayes, 1995). The hierarchical nature of stress also implies that different levels of stress have to be distinguished for different prosodic domains. Lexical stress determines the prominent syllable within a word. Beyond the lexical level, the prosodic phrase carries prominence on one of its lexical constituents, which may be realized as a pitch accent on the stressed syllable.

[2]Under certain circumstances, commas might serve as cues to prosodic phrasing (Steinhauer and Friederici, 2001; Steinhauer, 2003). The reliability of the comma-prosody correlation, however, crucially depends on the context (Chafe, 1988).

Likewise, among phrases within a sentence, one is assigned the nuclear accent, realized on the stressed syllable of the most prominent word within that phrase. The assignment of phrasal and sentence stress is mainly determined by the syntactic structure and the discourse context (Gussenhoven, 1983; Selkirk, 1995; Truckenbrodt, 2006). Word stress in German is lexically specified since it is not completely predictable from the segmental and syllabic structure of the word (Wiese, 2000).

Languages like German and English exhibit a general preference for an alternation of strong (i.e. stressed) and weak syllables, which manifests itself particularly in the avoidance of stress clashes, i.e. the avoidance of two adjacent syllables carrying main word stress. Stress clash avoidance has been demonstrated to affect language production in various ways: speakers might deviate from the citation form stress pattern to resolve a potential stress clash (Hayes, 1995; Selkirk, 1984); when faced with visually presented pseudo-words in sentential context, speakers have been shown to favor a stress pattern that maximizes rhythmic alternation (Kelly and Bock, 1988). The preference for rhythmic alternation may also have syntactic consequences for language production: given the choice, speakers preferably use syntactic constructions that prevent a stress clash (Anttila et al., 2010; Schlüter, 2005; Speyer, 2010).

As for auditory language comprehension, listeners were shown to be sensitive to rhythmic regularity in speech. Dilley and McAuley (2008) and Niebuhr (2009) report that listeners analyze the same lexically ambiguous syllable sequence differently depending on the linguistic rhythm (trochaic or iambic) established by the preceding context. Using event related potentials (ERP), Schmidt-Kassow and Kotz (2009) showed that listeners are sensitive to deviations from trochaic speech patterns when explicitly asked to judge the rhythmicity of the stimulus sentences. Niebuhr (2009) proposes that the phonetic rhythm has a guiding function in speech perception, in that it makes upcoming material predictable.

Warren et al. (1995) show that stress patterns on critical words have the potential to impinge on the syntactic analysis of temporarily ambiguous sentences. Specifically, their findings suggest that the perception of stress shift on critical words augments the cues to upcoming phrase boundaries even before such a boundary is encountered.

3.1.2 The generation of prosody in reading

Skilled readers produce prosody in accordance with the syntactic structure (Koriat et al., 2002) and also with the information structural analysis of the text. These factors especially influence accentuation and prosodic phrasing, implying that reading aloud simultaneously involves syntactic parsing, the interpretation of context, and the production of accordant prosody. The involvement of prosody in silent reading is less obvious, especially given the lack of a clear correlate of prosody in written text.

Recent research by Ashby and colleagues (Ashby and Clifton, 2005; Ashby and Martin, 2008; Ashby and Rayner, 2004) verifies the involvement of prosodic processing in silent reading on the lexical level. Ashby and Clifton (2005) demonstrate that readers fixate words with two stressed syllables (*situation*) longer than words with only one stressed syllable (*authority*), irrespective of the word length and frequency. Employing eye-tracking and ERP, Ashby and Martin (2008) find that readers routinely activate a prosodic phonological representation of the lexical items within the first 100 ms upon visual encounter. Ashby and Martin (2008) take this as evidence for an early speech-like phonological representation of the text being read.

The notion of speech-likeness suggests that the implicit prosody generated in the reading process is not to be understood as a simple concatenation of lexical prosodic structures. Instead, speech prosody is supralexical in nature, a condition that is evidenced, for example, by the stress shift phenomenon. To put it differently, if implicit prosody were speech-like, it should be subject to conditions of linguistic rhythm and the preference for an alternation of strong and weak syllables. Direct evidence for effects of linguistic rhythm in silent reading, however, is currently missing.

3.1.3 The role of implicit prosody in written sentence comprehension

Since prosody is not explicitly encoded in the graphemic string, its role in written sentence comprehension has been controversial: it is unclear whether the prosodic representation only reproduces the syntactic analysis by the reader (Kondo and Mazuka, 1996; Koriat et al., 2002) or whether implicit prosody itself contributes to the syntactic analysis during written sentence comprehension (Bader, 1998; Fodor, 1998, 2002).

A number of studies indicate that the silent prosody readers impose on the written text does affect the syntactic analysis. Bader (1998) finds that syntactically ambiguous sentences induce stronger processing difficulties in reading when the competing syntactic structures differ with respect to their prosodic features. He proposes the *Prosodic Constraint on Reanalysis* stating that revising a syntactic structure is particularly difficult if it necessitates a concomitant reanalysis of prosodic structure. Bader (1998) substantiates this proposal with reading data on temporarily ambiguous structures, the readings of which differ with respect to accent placement. Breen and Clifton (2011) examine the processing of lexical stress on noun-verb homographs (present – present) in syntactically ambiguous structures. Their results suggest that the reanalysis of lexical stress aggravates the resolution of syntactic ambiguities in silent reading.

Other studies focus on the effect of phrase length in relation to syntactic attachment preferences. Hirose (2003), Hwang and Schafer (2009) and Hwang and Steinhauer (2011) found that readers posit syntactic clause boundaries in temporarily ambiguous sentences based on the length of the preceding constituent. This leads to reading difficulties if the boundary turns out to be incompatible with the upcoming material. Several studies underpin the implicit prosodic effect in silent reading with consistent data obtained from oral reading experiments (Augurzky, 2006; Hirose, 2003; Hwang and Schafer, 2009; Jun, 2003). Others, however, fail to find the predicted correlation of attachment preference and overt prosodic pattern (Bergmann et al., 2008; Jun, 2010).

In summary, the research reviewed here clearly favors an account which grants implicit prosody a functional role in written sentence comprehension. As to the question of when and how exactly prosodic processes constrain the syntactic analysis in reading, the research on implicit prosody so far suggests that at least a partial syntactic analysis of the critical words and phrases is required in order for implicit prosody to show its effects on written sentence comprehension. Augurzky (2006) concludes from a thorough review and her own ERP data that "the parser initially leaves the prosodic analysis underspecified" (p. 206). Accordingly, prosodic effects on interpretation in silent reading would only occur in a very late processing stage.

Other studies emphasize the immediate nature of the prosodic effect. Recently, Hwang and Steinhauer (2011) presented ERP evidence suggesting that relatively long phrases trigger the insertion of implicit prosodic boundaries; apparently, the processing mechanism immediately interprets the implicit prosodic break to signal a syntactic phrase boundary. Note, however, that, in order to evaluate the length of phrases, the processing mechanism has to merge several words to form such phrases in the first place.

Correspondingly, existing research on implicit prosody is consistent with the idea that reading prosody depends on a partial syntactic analysis of the text. This generalization, however, might be due to the fact that the experiments mostly scrutinize effects of larger prosodic domains (prosodic phrasing, phrasal accentuation) on syntactic parsing. More local prosodic features like lexical stress and linguistic rhythm have, as yet, been largely disregarded. These factors may, however, more directly affect the assignment of syntactic structure. Clearly, different prosodic features might have different effects on the process of reading comprehension.

The work by Ashby and colleagues (Ashby and Clifton, 2005; Ashby and Martin, 2008) suggests that (lexical-)prosodic information such as syllable structure and stress pattern of the words is available to the processing mechanism from very early on in reading. It would be astonishing if it were not used immediately, especially since such information may be meaningful for the comprehension process (cf. Dilley and McAuley, 2008; Niebuhr, 2009; Warren et al., 1995). Under the assumption of a speech-like prosodic-phonological representation in reading, and given their immediate availability, stress and linguistic rhythm should exert their influence from the very beginning of the parsing process.

The following experiments are designed to put this hypothesis to a test and to show that even the earliest steps of syntactic parsing (i.e. the determination of the syntactic category of an ambiguous lexical item) may be guided by the implicit rhythm that emerges from the stress patterns readers impose on the written words.

3.2 Experiments

Given the general preference for the alternation of strong and weak syllables in German, it is predicted that a stress clash is avoided wherever more rhythmic alternatives are available. Despite the lack of explicit encoding of stress in written text, this should be true for reading aloud as well as for silent reading if readers indeed generate a speech-like phonological representation as proposed by Ashby and Martin (2008). This has consequences for the syntactic processing of the sentence: in the face of an ambiguous structure that involves a stress clash in one reading but not in the other, there should be a preference for the version without stress clash. This hypothesis will be tested in two reading experiments.

The object of investigation is syntactically ambiguous structures like (1), the two readings of which are differentiated prosodically by accentuation (stressed syllables underlined, accented syllables in capital letters).

(1) Der Polizist sagte, dass man...
The policeman said that one ...

a. ... nicht mehr NACHweisen kann, wer der Täter war. TEMP-INI
... couldn't prove anymore who the culprit was.

b. ... nicht mehr erMITteln kann, wer der Täter war. TEMP-MED
... couldn't determine anymore who the culprit was.

c. ... nicht MEHR nachweisen kann, als die Tatzeit. COMP-INI
... couldn't prove more than the date of the crime.

d. ... nicht MEHR ermitteln kann, als die Tatzeit. COMP-MED
... couldn't determine more than the date of the crime.

In (1), two different syntactic analyses of *mehr* are reflected in different prosodic renderings. In (1-a) and (1-b), *mehr* is part of the temporal adverbial *nicht mehr* (TEMP) and remains unaccented.[3] In this case, the following verb receives the main phrase accent. In (1-c) and (1-d), *mehr* is a comparative quantifier (COMP) that serves as a complement to the verb. In its function as complement to the verb, *mehr* receives main phrase accent, i.e. it is marked by a rising pitch accent. When preceded by an accented complement as in (1-c) and (1-d), the verb typically need not bear an accent (Truckenbrodt, 2006).

Since accent information is not encoded orthographically, the sentences are disambiguated in written text only after the verb complex, i.e. in the phrase that closes the sentence. In the temporal reading, the disambiguating phrase is a sentential argument of the verb that follows the ambiguous *mehr*.

In the comparative reading, *mehr* itself is the complement of the verb and the disambiguating phrase is the extraposed comparative complement of *mehr* introduced by the standard marker *als* (engl. *than*).[4]

A similar syntactic ambiguity involving *mehr* was studied by Bader (1996a) in a self-paced reading experiment. His results suggest that, generally, the temporal, unaccented reading of *mehr* is preferred over the comparative, accented reading, which Bader attributes to a general avoidance of (implicitly) accenting function words. Another possible source of this preference might be a higher frequency of the temporal *mehr*-construction.

For the purpose of this experiment, the rhythmic environment is systematically varied at the verb following *mehr*.[5] The verb has either initial stress (INI) as in (1-a) (TEMP-INI) and (1-c) (COMP-INI) or medial stress (MED) as in (1-b) (TEMP-MED) and (1-d) (COMP-MED). Condition COMP-INI, as opposed to all other conditions, involves a stress clash that is brought about by the adjacency of accented *mehr* and a verb with initial stress. The effect of this rhythmic imperfection on syntactic parsing will be tested in oral reading (experiment I) and in silent reading (experiment II).

[3] The semantics of the lexical unit *nicht mehr* in the temporal adverbial sense cannot be analyzed compositionally. It is therefore questionable whether the graphemic word *mehr* has an independent lexical status in this context.

[4] According to German comma rules, the sentential complement in (1-a) and (1-b) (temporal reading) is separated by a comma. As for the comparative reading, a comma is required only if the *als*-phrase is a clause, i.e. if it features an overt main verb. Although two-thirds of the comparative items in this experiment do not occur with a sentential disambiguating phrase but with an NP, the comma is set throughout to ensure comparability across conditions.

[5] Bader (1996) did not investigate rhythmic effects.

3.3 Experiment I

A speech production experiment was set up to test the influence of the rhythmic environment on the resolution of the local syntactic ambiguity concerning the word *mehr* in sentences like (1) in oral reading. The experiment consisted of two sessions in direct succession. In the first session, the 'unprepared session', participants read the stimuli out loud without advance preparation, i.e. without having knowledge of the disambiguation prior to executing the task. This way, the realization of accent on the critical word *mehr* should reflect the initial analysis unaffected by the disambiguating context. In the second session, the 'prepared session', participants were asked to familiarize themselves with the complete sentences before reading them out aloud.

3.3.1 Materials

24 sets of sentences like (1) were devised that contain a local syntactic ambiguity in writing but are unambiguous when spoken because of relevant prosodic cues. The actual sentences used in the experiments are listed in the appendix. All critical verbs following *mehr* are obligatorily transitive verbs that can take an NP or a sentential object to satisfy their argument structure requirements. The critical verbs are all trisyllabic, prefixed verbs that appear in their infinitival form and precede an inflected modal verb.

Validation of materials

The 12 verbs with initial stress and the 12 verbs with medial stress were matched with respect to word-form frequency and length. Word form frequencies were obtained from the Leipzig Wortschatz corpus (http://wort schatz.uni-leipzig.de/), which consists of approximately 50 million sentences of German newspaper text collected between 1994 and 2008. The mean logarithmized frequency is 7.14 (1.075 standard deviation) for verbs with initial stress and 6.93 (1.3) for verbs with medial stress. A linear model that evaluates the word frequency against the verb type does not suggest any significant difference between the two types (*F=0.186, df=22, p=0.67*).

Since the sentences with initial versus medial verb stress (examples in (1)) differ not only with respect to the verbal stress pattern but – necessarily – also with respect to the semantics of the verb (despite some effort to choose semantically similar verbs), a validation of the comparability of the conditions is required. To this end, all items were subjected to a sentence rating study.

The experimental sentences were distributed over four lists using a Latin square design with conditions counterbalanced across lists. In this way, each list presented

24 experimental sentences, six from each condition. In each list, the experimental sentences were interspersed with 76 filler sentences from four unrelated experiments. The order of the items was pseudo-randomized using the Mix randomization tool by van Casteren and Davis (2006) such that items from the same experiment had a minimum distance of three and items from the same experimental condition had a minimum distance of six. Each list was printed on A4 paper in landscape layout with each sentence presented on a single line.

Forty-six first-year undergraduate students from the University of Potsdam, all naïve to the purpose of the experiment, took part in the rating study for course credit or payment. They were each given one of the four lists. The subject's task was to rate every sentence on a seven point Likert scale (1 - easy and perfectly acceptable sentence – 7 - incomprehensible, unacceptable sentence) and note the respective number next to each sentence on the sheet. No time constraints were given. All participants completed the rating task within 40 minutes.

Of the total 1104 sentences, 48 (4%) had missing or unidentifiable ratings. The 1056 obtained ratings were treated as numerical values. The boxplot in Figure 1 depicts the median and the distribution of the ratings by condition.

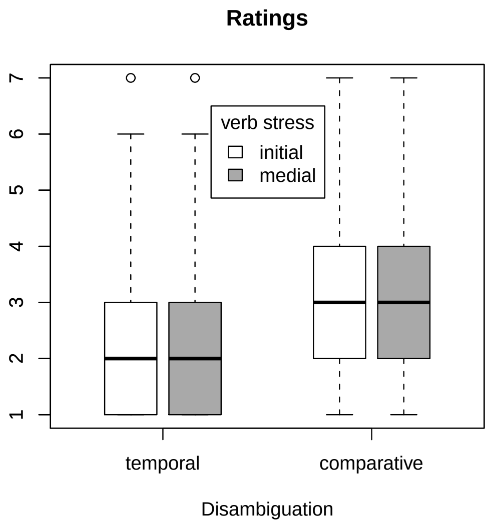


Figure 3.1: Boxplots representing the distribution of ratings broken down by 'disambiguation' and 'verb stress'.

The ratings were evaluated against the crossed fixed factors 'disambiguation' (with the two levels TEMP and COMP) and 'verb stress' (with the two levels INI and MED) using a linear mixed model (Bates and Sarkar, 2007; Gelman and Hill, 2007). Participant and item were treated as random variables. Table 1 summarizes the results of the model.[6] The model reveals a significant main effect for 'disambiguation'. The effect for 'verb stress' and the interaction are non-significant. The rating results thus do not indicate any difference in terms of acceptability of the sentences that is systematically attributable to the implicit rhythmic environment brought about by the stress pattern of the verb. However, the significant main effect of 'disambiguation' shows that the temporal reading of *mehr* (TEMP) is on the whole more acceptable than the comparative versions (COMP).

Coefficient	Estimate	Std. Error	t-value
disambiguation	0.2696	0.0374	7.207
verb stress	0.0077	0.0375	0.206
disamb × v-stress	0.0392	0.0374	1.048

Table 3.1: Results of linear mixed model evaluating the ratings against the crossed fixed factors 'disambiguation' and 'verb stress'.

3.3.2 Experimental procedure

For the oral reading experiment, the experimental sentences were again distributed over four lists with conditions counterbalanced across lists. In each list, the 24 experimental sentences were embedded in 69 filler sentences from four unrelated experiments. The total of 93 items was pseudo-randomized for each subject independently, such that sentences from the same experiment had a minimum distance of three items and sentences from the same experimental condition had a minimum distance of eight items. Participants saw the same list of items in the same order in both unprepared and prepared sessions of the experiment.

The experiment took place in an anechoic room with an AT4033a Audio-Technica studio microphone. Each participant was seated in front of a 15" computer screen with the microphone placed approximately 30 cm from the participant's mouth. A keyboard was placed on a table in front of the subject. Recordings were made on a computer using the RecordVocal function of the DMDX (Forster and Forster, 2003) and a C-Media Wave sound card at a sampling rate of 44.1 kHz with 16 bit resolution.

Each of the two sessions was preceded by three example stimuli (not related to any of the experimental stimuli) for the participants to familiarize themselves with the task.

For the unprepared session, the DMDX presentation was programmed for each item as follows: First, only the first one or two words (the sentence initial subject

[6]Since, in linear mixed models, determining the precise degrees of freedom is non-trivial, the t-values are approximations. An absolute t-value of 2 or greater indicates statistical significance at $\alpha = 0.05$.

noun phrase or proper name) were presented on the screen. Participants were told to familiarize themselves briefly with these words. They were instructed to then press the space bar, inducing the presentation of the entire sentence. Participants were asked to start reading the sentence out loud immediately as it appeared on the screen and to do so as fluently as possible. Pressing the space bar automatically initiated the recording. After a fixed recording time of five seconds, the procedure was repeated for the next item. For each sentence, only one realization per subject was recorded. No corrections were recorded in the case of hesitations or slips of the tongue.

After completion of the unprepared reading session, participants were encouraged to take a short break of approximately five minutes, which was followed by the prepared session. The item presentation differed from the unprepared session in that readers were presented with the whole sentence from the start and were told to familiarize themselves with the sentence before reading it out loud. Again, when ready to read out loud, readers were asked to press the space bar to initiate the recording. This time, pressing the space bar did not change the visual presentation. For each item, the recording time was set to five seconds, after which the next item appeared on screen. For each item, again only one realization per subject was recorded.

Participants

Twenty-four female first-year undergraduate students from the University of Potsdam took part in the experiment. All were native speakers of German and naïve as to the purpose of the experiment. They either received course credit or were paid 5 Euros for their participation.

3.3.3 Predictions

In unprepared reading, i.e. if readers are unaware of the disambiguation, accentuation of *mehr* should be avoided given that the unaccented, temporal reading is processed more easily and is generally preferred (c.f. Bader (1996a) and section 3.1.1.). Moreover, the predicted avoidance of stress clash is hypothesized to lead to a higher number of unaccented realizations of *mehr* in the case of a following verb with initial stress.

On the assumption that successful reading necessitates conformity of prosodic and syntactic structure, realizations of *mehr* that are prosodically incompatible with the disambiguating region should lead to reading difficulties – that is, readers might be led down the garden path if their prosodic realization of *mehr* turns out to be infelicitous. Since reading comprehension is strongly correlated with reading fluency, the difficulties should manifest themselves in hesitations or a slowdown in speech (e.g. Fuchs et al., 2001) once the reader reaches the disambiguating region.

As for prepared reading, the disambiguation is known to the reader before oral realization. The disambiguation reveals the lexical-syntactic status and with it the appropriate accent for *mehr* – it should thus be the decisive factor for the accentuation of *mehr*. The immediate rhythmic environment does not alter the grammatical requirement of accentuation on *mehr* and therefore should not have a systematic effect.

3.3.4 Data analysis

All in all, 1152 experimental sentences were recorded, 576 in the unprepared session and 576 in the prepared session. The sentences from the two sessions were independently judged by two students each. The judges were not informed about the conditions and the purpose of the experiment before completion of their job. Their task was i) to note slips of the tongue and disfluencies in the part of the sentence up to but excluding the disambiguating phrase, and ii) to determine for each sentence if the word *mehr* was accented or not, i.e. if it was to be understood as a comparative complement or as a temporal adverbial. In order to avoid an influence of the disambiguating region on the judgments, all sound files were cut after the verb complex prior to the judgment process. The sentences were presented to the judges in randomized order. The judges were paid for their work.

For ease and clarity of exposition, the results of the prepared reading task will be reported before the results of the unprepared reading session.

3.3.5 Results

Results for prepared reading

Twenty-four (4%) of the total 576 sentences were marked by at least one of the judges as non-fluent or containing slips of the tongue in the region preceding the disambiguating phrase. A generalized linear mixed model (GLMM) with a binomial link function (Bates and Sarkar, 2007; Gelman and Hill, 2007; Quené and van den Bergh, 2004) was fitted to check whether the distribution of flawed sentences is related to the experimental factors. The fixed factors of this model were i) 'disambiguation' (COMP vs. TEMP) and ii) 'verb stress' (INI vs. MED) with flawed versus fluent realization as the dependent variable; participant and item were included as random effects (grouping variables). Orthogonal contrast coding was applied (factor 'disambiguation': comparative=1, temporal=-1; factor 'verb stress': initial=1, medial=-1). This model does not reveal any systematic influence of the controlled factors 'disambiguation' and 'verb stress' or their interaction on the distribution of flawed sentences (all z-values <|2|, all p-values >.2). As for the 552 fluent sentences, the assessments of the two judges concerning the accentuation of *mehr* concur in 532 cases (97%). The bar plot in Figure 2 shows the percentage of accented *mehr* by condition for the consistently judged sentences. The target word was perceived as accented in the comparative readings (COMP-INI, COMP-MED) in around 90% of the cases; as for the temporal reading (TEMP-INI, TEMP-MED), *mehr* was perceived as accented in less than 10% of the cases.

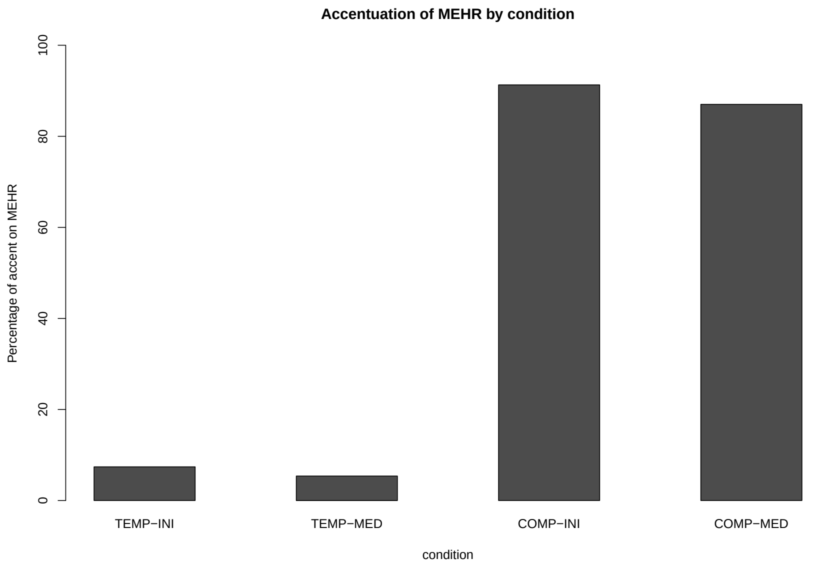


Figure 3.2: Accentuation of *mehr* as determined by judges broken down by condition in prepared reading experiment.

The accentuation status of *mehr* was again evaluated with a GLMM incorporating the same fixed factors and grouping variables as above. In line with the above predictions, this model confirms a single significant main effect for the fixed factor 'disambiguation'. The main effect for 'verb stress' and the interaction remain non-significant (cf. Table 2).

Coefficient	Estimate	Std. Error	z-value	p-value
disambiguation	2.96412	0.20989	14.122	<0.001
verb stress	0.26280	0.18773	1.400	0.162
disamb × v-stress	0.04809	0.18752	0.256	0.798

Table 3.2: Results of GLMM on accentuation of *mehr* in consistently judged sentences in prepared reading experiment.

Results for unprepared reading

In the unprepared session, 63 sentences (11%) were non-fluent or contained slips of the tongue in the region preceding the disambiguating phrase, as determined by at least one of the judges. As for the accentuation status of *mehr*, the judges agreed on 495 of the 513 fluent sentences (96%). The 495 consistently judged sentences were hand-annotated by a phonetically trained student who was blind to the purpose of the experiment and to the judgments of her fellow students. For each of the 495 sentences, the critical region starting with *nicht* up to the end of the verb complex was segmented into words and syllables and labeled accordingly.

Flawed sentences
The number of flawed sentences is relatively high (n=63, 11%), which can be partly explained by the task (unprepared reading) and the length of the sentences (10 words up to the disambiguating region). It was checked whether the distribution of flawed sentences is systematically related to the controlled factors of the experiment. No significant effect was found for either of the fixed factors ('disambiguation': *z=-0.501, p=0.62*; 'verb stress': *z=-0.747, p=0.46*), or the interaction (*z=1.017, p=0.31*), suggesting that the controlled variables do not systematically influence the distribution of flawed sentences.

Judgments on realizations of mehr
The bar plot in Figure 3 displays the percentages of accented *mehr* as perceived by the judges in the four conditions. In total, *mehr* was perceived as accented in about 24% of the cases.

Speakers accented *mehr* in 20% of the sentences with comparative disambiguation. Twenty-seven percent of the occurences of *mehr* were judged as accented in sentences with temporal reading. When the verb following *mehr* has medial stress, speakers accented *mehr* in 28% of the sentences, compared to 19% when the verb has initial stress.

A GLMM was fitted with perceived accentuation of *mehr* as the dependent variable. The fixed factors of this model are again i) 'disambiguation' (TEMP vs. COMP) and ii) 'verb stress' (INI vs. MED). Speakers and items served as random effects. This model (Table 3) yields a significant main effect for the stress position on the verb. The effect of the disambiguating region is significant, too. The interaction of stress position and disambiguation is not significant.

A comparison with the prepared reading data reveals that the accentuation status of *mehr* is frequently inappropriate relative to the subsequent disambiguation. In conditions COMP-INI and COMP-MED in particular, only 20% of the trials were realized with the required accent on *mehr*. In contrast, *mehr* congruously

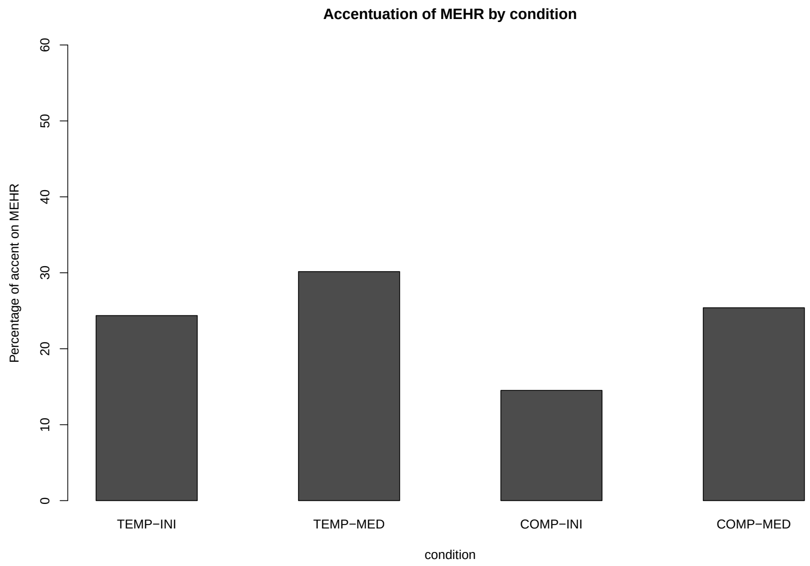


Figure 3.3: Accentuation of *mehr* as determined by judges broken down by condition in unprepared reading experiment.

Coefficient	Estimate	Std. Error	z-value	p-value
disambiguation	-0.2613	0.1258	-2.077	0.038
verb stress	-0.3601	0.1262	-2.853	0.004
disamb × v-stress	-0.1351	0.1254	-1.077	0.282

Table 3.3: Results of GLMM on perceived accentuation of *mehr* in consistently judged sentences in unprepared reading experiment.

remained unaccented in the temporal conditions TEMP-INI and TEMP-MED in 72% of the cases. Given the abovementioned main effect of verb stress on the realization of accent, the avoidance of accent on *mehr* in the comparative conditions should result in even more instances of realizations that are incompatible with the disambiguating region when the verb features initial stress. To check for this interaction, a GLMM was fit. The dependent variable this time was the appropriateness of accentuation relative to the disambiguating region. The model confirms a clear main effect of 'disambiguation' and reveals that the interaction between 'disambiguation' and 'verb stress' is significant (cf. Table 4).

Phonetic analysis of accented vs. unaccented realizations

Overall, perceived accentuations of the target word are conspicuously rarer in unprepared reading as compared to prepared reading. This is most likely due to the general preference for the unaccented, temporal reading that was attested by Bader (1996) and confirmed in the sentence rating study above.

Coefficient	Estimate	Std. Error	z-value	p-value
disambiguation	-1.2057	0.1090	-11.060	<0.001
verb stress	-0.1007	0.1090	-0.924	0.3557
disamb × v-stress	-0.2471	0.1090	-2.267	0.0234

Table 3.4: Results of GLMM evaluating the compatibility of accentuation of *mehr* relative to the disambiguating region.

In order to exclude misperception by the judges, their assessment was validated by means of a phonetic analysis. Also, since listeners may perceive prominence patterns on syllable sequences in context even in the absence of definite acoustic cues for such a pattern (Dilley and McAuley, 2008), a validation of their judgments is appropriate. Hence, the syllable durations and pitch contours of sentences with perceived accented and unaccented *mehr* were compared. Specifically, the region starting with *nicht* up to the modal verb preceding the disambiguating phrase was analyzed.

The upper panel of Figure 4 shows the grand average pitch contours in the critical region broken down by accentuation of *mehr* and the stress position on the following verb. The pitch contours were created by dividing each syllable in the region of interest into three equal-sized intervals and interpolating the normalized mean F0 for each of these intervals; the normalization factor used is the inverse of the maximum F0 of each sentence. The bar plots in the lower panel display the respective average syllable durations in milliseconds. Clearly, the tokens of *mehr* that were perceived as accented display longer durations compared to unaccented tokens. Moreover, there is a clear rising pitch contour on *mehr* in the accented versions (black lines), indicating the realization of a pitch accent on this word. The versions with unaccented *mehr* (grey lines) show falling pitch on the critical word and the rise appears only later on the stressed syllable of the following verb, which, in these cases, carries the phrase accent. The accentuation of *mehr* appears to already have small effects on the duration and pitch contour of the preceding *nicht* and continues to have durational effects on the realization of the following verb. Irrespective of accentuation on *mehr*, the modal verb ends on a relatively high pitch, indicating a continuation rise preceding the disambiguating phrase.

A linear mixed model with subject and item as random effects confirms a significant effect of perceived accentuation on the prosodic rendering of *mehr*. The dependent variable of this model is the pitch slope on *mehr*, i.e. the difference between the F0 values at the onset and the offset of *mehr* divided by the duration of *mehr*. The perceived accentuation serves as the fixed effect, yielding a coefficient estimate of 66.36 with a standard error of 4.88 (*t-value= 13.60*). The phonetic analyses confirm the prosodic difference between perceived accented and unaccented versions and thus validate the judgments.

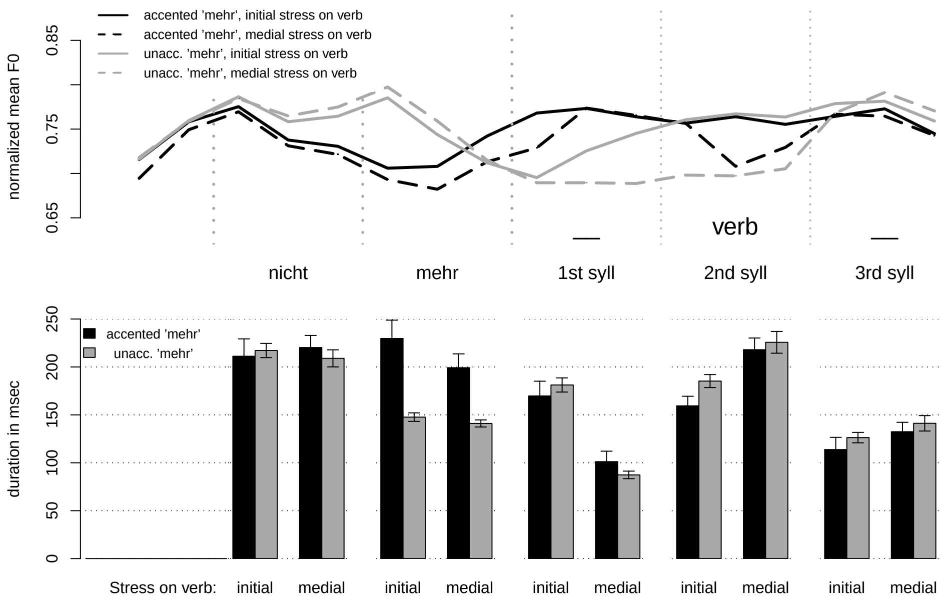


Figure 3.4: Grand average pitch contours (upper panel) and durations (lower panel) for each syllable in the region starting with *nicht* up to the critical verb broken down by realization of accent and verbal stress pattern.

Phonetic analysis of garden path effect

On the assumption that the realization of accent on *mehr* conforms to the syntactic analysis, the readers/speakers should experience comprehension difficulties when the realization of *mehr* is incompatible with the disambiguating region. In fluent oral reading, the reader's eyes are a few words ahead of the voice; hence, the slowdown in speech should already be observable at the beginning or even before the disambiguating region is spoken aloud. Correspondingly, the modal verb and the pause preceding the disambiguating phrase might be affected by the slowdown and show longer durations when the disambiguating region is inappropriate relative to the accentuation applied on *mehr*. To test for this garden path effect, the durations of both the modal verb and the pause preceding the disambiguating phrase were summed and evaluated. Specifically, the duration from the onset of the modal verb up to the onset of the disambiguating phrase was measured. By inclusion of the modal verb, effects of final lengthening due to the clause break are included in the analysis.

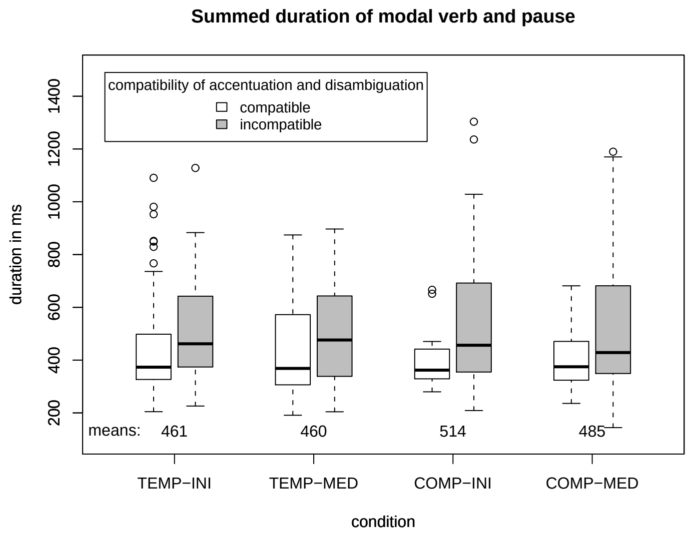


Figure 3.5: Duration of modal verb and following pause before the disambiguating clause in the four conditions broken down by compatibility of accentuation.

The boxplots in Figure 5 depict the distribution of the clause break durations (modal verb plus pause) in each condition broken down by the compatibility of the accentuation applied by the reader relative to the disambiguation. When the speakers' realization of *mehr* is inappropriate relative to the disambiguating phrase (incompatible realizations), the duration data indicates a marked slow-down compared to the appropriate realizations.

A linear mixed model with the fixed factors 'compatibility of realization', 'disambiguation', and 'verb stress' and subjects and items as random effects yields a significant main effect for 'compatibility' on duration, confirming that speakers indeed slow down when the disambiguating region does not conform to the accentuation of the ambiguous *mehr*. The effect of 'disambiguation' is not significant, nor is the effect of 'verb stress' or any of the interactions. The parameters of this model (cf. Table 5) thus suggest that the compatibility of accentuation has similar effects irrespective of the presented condition.

Note, however, that some of the factors of the model are highly correlated: due to the preference for the unaccented temporal reading of *mehr*, significantly more compatible realizations were made in conditions TEMP-INI and TEMP-MED as compared to COMP-INI and COMP-MED; moreover, more compatible realizations were produced in condition COMP-MED compared to the clash condition COMP-INI. A second analysis (cf. Table 6) evaluates the duration data against the fixed factors 'disambiguation' and 'verb stress' only, thereby avoiding any correlation. This model reveals a significant main effect of 'disambiguation'.

Coefficient	Estimate	Std. Error	t-value
compatibility	-0.0901	0.0164	-5.48
disambiguation	-0.0094	0.0162	-0.58
verb stress	0.0078	0.0162	0.48
comp × disamb	-0.0208	0.0188	-1.11
comp × v-stress	-0.0175	0.0164	-1.07
disamb × v-stress	-0.0017	0.0164	-0.11
comp × disamb × v-stress	-0.0127	0.0165	-0.77

Table 3.5: Results of linear mixed model evaluating the summed duration of modal verb and pause against compatibility of accentuation, disambiguation and verb stress.

Coefficient	Estimate	Std. Error	t-value
disambiguation	0.0412	0.0137	3.02
verb stress	0.0197	0.0137	1.45
disamb × v-stress	0.0166	0.0137	1.21

Table 3.6: Results of linear mixed model evaluating the summed duration of modal verb and pause against the factors disambiguation and verb stress.

Although the interaction does not reveal a significant effect, closer inspection shows that the stress clash condition COMP-INI leads to significantly longer durations compared to COMP-MED (*coeff. estimate=0.0355, std.err.=0.0169, t-value=2.10*). In contrast, the durational difference between the two temporal conditions TEMP-INI and TEMP-MED does not appear to be systematic (*coeff. estimate=-0.0005, std.err.=0.0203, t-value=0.02*).

In summary, the analysis concerning the duration of the modal verb plus the pause preceding the disambiguating region presents evidence that incompatible realizations of accent on *mehr* lead to a garden path effect, indicating that speakers made a syntactic commitment when choosing the accent status of *mehr*. On average, durations of the region of interest were longer in the comparative reading (COMP-INI and COMP-MED) compared to the temporal versions (TEMP-INI and TEMP-MED), with the longest durations found in the clash condition COMP-INI. Since the increase in duration due to inappropriate accentuation appears to be similar across conditions (cf. Figure 5 and Table 5), the number of compatible vs. incompatible accentuations in the four conditions is most likely responsible for the difference in mean duration between conditions.

3.3.6 Discussion

The accentuation patterns of the target word in the prepared reading session conform to expectations: the ambiguous item *mehr* is accented when used as a comparative (engl. *more*) but remains unaccented in the temporal reading of *nicht mehr* (engl. *not anymore, no longer*). That is, if readers have full access to the disambiguating material before starting to read out loud, they audibly use accentuation to signal the appropriate variant of *mehr*. The immediate rhythmic environment (the verb stress manipulation) does not systematically contribute to the accentuation status of *mehr* in the prepared reading session. This also fits the expectations according to which the requirement for the accentuation of *mehr* is solely driven by its syntactic status.

As for unprepared reading, readers chose to accent the critical word *mehr* in just under 25% of the cases, indicating a preference for the unaccented, temporal version. This effect conforms to the findings by Bader (1996a) and the rating study (3.1.1), which showed higher acceptability of the temporal conditions TEMP-INI and TEMP-MED as opposed to conditions COMP-INI and COMP-MED. Importantly, the judgments concerning the accentuation of *mehr* in unprepared reading reveal a significant main effect for the verbal stress pattern on the realization of *mehr*, confirming that the accentuation of the target word is systematically influenced by the immediate rhythmic environment: as hypothesized, speakers avoid accenting *mehr* when this would induce a stress clash configuration with the following verb. As predicted, this rhythm-induced avoidance of accent leads to a significantly higher number of inappropriate realizations in the context of the comparative disambiguating region.

Unexpectedly, accent on *mehr* was realized significantly more often in temporal versions (TEMP-INI and TEMP-MED), i.e. when the disambiguating region requires *mehr* to remain unaccented (cf. the significant main effect of 'disambiguation' in Table 3). This effect seems to suggest that the readers used information in the disambiguating phrase for the assignment of accent on *mehr*, but it remains unclear what type of information this might be and what makes this information misleading. In any case, this effect shows that the disambiguating material does not have a facilitating effect on the appropriate realization of *mehr* in unprepared reading. A comparison of the accentuation patterns in unprepared reading with those of the prepared session indicates that the readers were most likely unaware of the disambiguating information in the unprepared session. Also, the manifestation of the verb stress effect in unprepared reading suggests that readers use implicit rhythmic cues more readily than whatever information they have about the disambiguating phrase when determining the accentuation status of *mehr* in this task.

The phonetic analysis of the accented and unaccented versions of *mehr* confirms the validity of the judges' perceptions. As expected, accented *mehr* is realized with a strong rise in pitch and longer duration compared to unaccented versions. The duration data at the clause break provide an indication that the accentuation involves a syntactic commitment on the part of the speakers. The region before the disambiguating clause is significantly prolonged when realizations of *mehr* are incompatible with the disambiguating region. This slowdown is indicative of a garden path effect. The data suggest that the readers/speakers in fact assign syntactic features to *mehr* according to their realization of accent on this item and experience integration difficulty if the disambiguating region does not conform to the prior prosodic realization.

Overall, the first experiment confirms that reading prosody is dependent not only on the syntactic structure and the lexically determined syllable and stress information of the words in the written string, but also on the supralexical linguistic rhythm emerging from the concatenation of single words. Specifically, the experiment presents evidence for the hypothesis that rhythmic expectancy, i.e. the avoidance of stress clashes, affects the prosodic realization and, consequently, syntactic parsing in unprepared oral reading. Beyond a general preference for the unaccented temporal reading of *mehr*, the local rhythmic environment demonstrably constrains the respective assignment of the syntactic features. That is, if the syntactic structure is underspecified, the reader chooses the accentuation and, consequently, the syntactic analysis that best conforms to syntactic and prosodic well-formedness constraints. This interpretation of the results implies that readers evaluate the syntactic structure of written material as a function of the prosodic environment which is generated by a process of phonological recoding. At first glance, this idea is at odds with existing research on reading prosody that emphasizes the dependence of prosody on the syntactic analysis (Kondo and Mazuka, 1996; Koriat et al., 2002). Those experiments on reading prosody, however, are chiefly concerned with syntactically unambiguous structures and focus on the relation of larger syntactic constituents and prosodic phrasing. More local prosodic features like stress and linguistic rhythm may therefore affect the assignment of syntactic structure in the ambiguous region without contradicting research on the relation of syntax and prosodic phrasing. As mentioned above, different prosodic features might have different repercussions at different processing stages in reading comprehension.

In any case, experiment I does not allow firm conclusions to be drawn about the precise relation of prosodic and syntactic processes in reading. The dependent measures evaluated so far are bound to speech production in oral reading, which is known to lag behind sentence comprehension (Levin and Addis, 1979). It can thus only indirectly inform us about the interplay of syntax and prosody in the comprehension processes. Moreover, while oral reading necessarily involves reading prosody, the involvement of prosody in silent reading is less evident. Data that is arguably more time sensitive and therefore more informative about the role of implicit prosody in written sentence comprehension comes from the relevant sense organ, i.e. the eye-movement record (Rayner, 1998; Clifton et al., 2007).

3.4 Experiment II

The notion of a speechlike phonological representation in silent reading implies that readers have rhythmic expectancies. They should especially avoid representations of adjacent stressed syllables whenever more rhythmic alternatives are accessible. The present experiment tests this hypothesis using the same material as in experiment I applying eye-tracking methodology for silent reading. The example sentences are repeated in (2).

(2) Der Polizist sagte, dass man...
The policeman said that one ...

a. ... nicht mehr NACHweisen kann, wer der Täter war. TEMP-INI
... couldn't prove anymore who the culprit was.

b. ... nicht mehr erMITteln kann, wer der Täter war. TEMP-MED
... couldn't determine anymore who the culprit was.

c. ... nicht MEHR nachweisen kann, als die Tatzeit. COMP-INI
... couldn't prove more than the date of the crime.

d. ... nicht MEHR ermitteln kann, als die Tatzeit. COMP-MED
... couldn't determine more than the date of the crime.

As in experiment I, we hypothesize that readers should choose the syntactic category of ambiguous words in such a way as to accord with rhythmic preferences. In the case of the structures in (2), *mehr* should be computed as an unaccented temporal adverbial more often when followed by a verb with initial stress to avoid a stress clash configuration. This in turn should lead to increased reading difficulties in the disambiguating region if the comparative reading of *mehr* is required. That is, reading the disambiguating clause in the clash condition COMP-INI should be associated with higher processing costs compared to reading the rhythmically alternating COMP-MED. No such difference is expected between TEMP-INI and TEMP-MED as neither of them violates rhythmic preferences. Therefore, an interaction between the factors 'disambiguation' and 'verb stress' is predicted. Beyond this interaction, the clear preference for the unaccented version of *mehr*, which was attested in experiment I, should lead to increased reading difficulties in conditions COMP-INI and COMP-MED as compared to TEMP-INI and TEMP-MED.[7]

To test the influence of the rhythmic environment on the resolution of a local syntactic ambiguity in silent reading, eye-tracking methodology was employed. This method involves monitoring readers' eye-movements as they scan written text on a screen. The difficulty of identifying and integrating a given word is strongly correlated with the fixation patterns on and around that word. Tracking these patterns with a high temporal and spatial resolution therefore allows sentence comprehension processes to be studied in real time (Rayner, 1998). Syntactic parsing difficulties in the disambiguating region should be reflected in more fixations, longer fixation durations, and a higher probability of regressions in that area (Clifton et al., 2007).

3.4.1 Materials and methods

Stimuli

The set of 24 experimental items from experiment I was used for the eye-tracking experiment as well. Again, the factors 'disambiguation' and 'verb stress' were crossed in a 2×2 design to define the four conditions.

Participants

Forty-eight undergraduate students of the University of Potsdam took part in the experiment for course credit or were paid for participation. All of them reported normal or corrected-to-normal vision. None of them participated in experiment I.

Experimental procedure

The participants were seated in front of an IView-X eye-tracker (SensoMotoric Instruments) running at 240 Hz sampling rate and with 0.025 degree resolution. To ensure stability of the eye position, participants placed their heads in a frame with a chin rest. A camera within in the frame monitored the pupil of the participant's right eye during the entire experiment. Each sentence was presented on a single

[7] In the unprepared oral reading experiment, more instances of *mehr* were accented in the TEMP-conditions than in the COMP-conditions; this unexpected effect, however, does not alter the predictions for the silent reading experiment; the amount of appropriate realizations of *mehr* relative to the disambiguating region was still markedly higher in the TEMP-conditions (72%) than in the COMP-conditions (20%), confirming the preference for the unaccented, temporal reading.

line on a 17" monitor with 1024x768 pixel resolution. Stimulus presentation and recording of the eye-movements were controlled by Presentation software. The experimental sentences were divided into four lists such that experimental sentences and conditions were counterbalanced across lists and participants saw at most one sentence from each of the 24 item sets. Each participant was assigned one of the four lists, each of which contained 24 target items together with 76 filler sentences from four unrelated experiments in pseudorandomized order. A calibration procedure preceded the experiment: participants looked at 13 fixation points that appeared in random order to allow for gauging of the gaze position. This procedure was repeated after every 10–15 trials or when measurement accuracy was poor. To direct the participants' eyes to the beginning of the sentence, a fixation point was shown at the position of the leftmost character immediately before presentation of the trial. Directly upon fixation of this target, the sentence was displayed. Participants were asked to silently read the sentence and click a mouse button when finished. A forced choice comprehension question followed each trial, for example *'Hat der Polizist etwas gesagt?'* (Engl.: 'Did the policeman say something?'). Answering the question by mouse click triggered the presentation of the next item.

Defining the dependent measures and the regions of interest

Four eye-tracking measures that are considered standard measures in the literature on eye-tracking in sentence comprehension research (Rayner, 1998; Clifton et al., 2007) were used as dependent variables. These are i) first-pass reading time (FPRT), i.e. the sum of all fixation durations within a region until leaving the region, given that the region was fixated at least once; ii) second pass or re-reading time (RRT), that is the summed fixation time on a given region after first pass (including zero times if the region was not refixated); iii) the total fixation time (TFT); iv) the probability of regressing out of a region during first pass (RegrP), i.e. before material to the right of the region was fixated. In addition to these standard measures, the probability of skipping a word (SKIP) during first pass and the re-reading probability (RRP), i.e. the probability of refixating a region after first pass, were calculated. All dependent measures are examined on individual words in or near the disambiguating region.

FPRT and RegrP are assumed to reflect so-called 'early' processing stages and may indicate the difficulty associated with higher level lexical processing such as integrating words with the preceding context (Clifton et al., 2007). The skipping probability (SKIP) may reflect even earlier processes since the decision to fixate or skip a word during first pass is necessarily made on a word preceding the affected target word. The 'late' measures (RRT, RRP and TFT) are generally considered to reflect more general comprehension difficulties (Clifton et al., 2007).

Unfortunately, it is far from clear how to distinguish between integration difficulty on the one hand and general comprehension difficulty on the other. The interpretation of the dependent measures and the distinction between 'early' and 'late' processing stages depend on various factors such as type of ambiguity, strength of interpretation bias, type of disambiguation (morphological, syntactic, semantic) and also on the size of the region under examination. Therefore, the precise cognitive processes responsible for a particular dependent measure remain a matter of debate in eye-movement research in reading. In general, however, longer reading times and higher regression rates are associated with higher cognitive demand, while shorter fixation times and more frequent skipping may signal relative reading ease (Clifton et al., 2007; Rayner, 1998).

In this experiment, two words (full verb and modal verb comprising 12 to 18 characters altogether) intervene between the ambiguous word and the disambiguating region. Moreover, a clause boundary precedes the disambiguating phrase. Reading difficulties are expected to show up in the disambiguating phrase that follows the modal verb. The first word of this phrase disambiguates the temporal and comparative reading: for the temporal reading, it is either a (wh-)pronoun or a complementizer that introduces a sentential complement to the preceding verb complex (cf. TEMP-INI and TEMP-MED in (2)). In the comparative reading, the standard marker *als* introduces the comparative complement of *mehr*, which is the standard of comparison (cf. COMP-INI and COMP-MED in (2)). Since the first word of the disambiguating phrase is relatively short (3 or 4 characters), measures were analyzed on both the first (Region 1) and the second word (Region 2) of the disambiguating phrase. Additionally, fixation patterns on the modal verb that precedes the disambiguating region (Region 0) were analyzed in order to check for parafoveal effects (Kennedy and Pynte, 2005; Kliegl et al., 2006). This is motivated by the fact that Region 1 was frequently skipped (see below) and skipping of a word usually correlates with longer reading times on the preceding word (Kliegl and Engbert, 2005). Finally, to gauge a possible spillover effect, the eye-movement record of the last word of the disambiguating phrase (Region 3) was examined, too.

Data analysis

Due to miscalibrations, data from one participant was excluded from further analysis (only 5% of the subject's fixations were recorded as fixations on words).

Question response accuracies were computed. Only those trials that were responded to correctly and in which the critical verb following *mehr* was fixated during first pass were included in the statistical analysis of the eye-tracking measures. The *em* package by Logačev and Vasishth (2006) was used to calculate the dependent measures. For the statistics on FPRT and TFT, fixations shorter than 50 ms were removed and treated as missing values. In order to adjust for the skew in the data, fixation durations were log-transformed for inferential statistics (Gelman and Hill, 2007).

Statistical analysis
The fixation durations (FPRT, RRT, TFT) were analyzed using linear mixed models; skipping probability (SKIP), re-reading probability (RRP) and regression probability (RegrP) were modeled using generalized linear mixed models (GLMM) with binomial link function (Bates and Sarkar, 2007). As in experiment I, the dependent measures were evaluated against the factors 'disambiguation' (TEMP vs. COMP) and 'verb stress' (INI vs. MED) and the respective interaction. Participants and items were included as crossed random effects. Again, contrast coding was applied as in experiment I (factor 'disambiguation': comparative=1, temporal=-1; factor 'verb stress': initial=1, medial=-1).

3.4.2 Results

Various reading measures in several regions of the disambiguating phrase reveal an increase of processing costs in the comparative conditions relative to the temporal conditions, with most difficulty arising in the clash condition COMP-INI. Results for each region of interest will be detailed below.

Response accuracy

On average, participants answered 86% of the comprehension questions correctly. A GLMM that evaluates the error rates against the experimentally controlled factors does not reveal any significant influence of the fixed factors on the distribution of the erroneous answers (effect of 'disambiguation': *z-value=-0.474, p=0.64*; effect of 'verb stress': *z-value: 1.567, p=0.12*; interaction: *z-value=1.175, p=0.24*).

Reading measures

The reading measures for the Regions 0 through 3 are tabulated in Table 7. The results of all regions are discussed in the following. Inferential statistics for Regions 2 and 3 are tabulated in Tables 8 and 9 respectively.

Region 0: Word preceding the disambiguating phrase
The word preceding the disambiguating phrase is a mono- or disyllabic modal verb comprising 4 to 7 characters. The clash condition COMP-INI displays the highest FPRT, RRT and TFT in this region. Apparently, the values for conditions

Table 3.7: Raw reading measures (means) broken down by condition and region of interest.

		Region of interest			
Measure	condition	0	1	2	3
SKIP	TEMP-INI	0.15	0.45	0.24	0.2
	TEMP-MED	0.09	0.41	0.23	0.19
	COMP-INI	0.10	0.47	0.24	0.17
	COMP-MED	0.08	0.49	0.33	0.21
FPRT (Std.Err.) in ms	TEMP-INI	224 (8)	248 (10)	238 (10)	314 (18)
	TEMP-MED	222 (8)	233 (9)	236 (9)	305 (20)
	COMP-INI	245 (10)	229 (9)	231(9)	351(19)
	COMP-MED	236 (10)	216 (9)	240 (10)	313 (17)
RegrP	TEMP-INI	0.21	0.13	0.22	0.55
	TEMP-MED	0.25	0.10	0.28	0.54
	COMP-INI	0.15	0.07	0.22	0.68
	COMP-MED	0.17	0.07	0.23	0.55
RRP	TEMP-INI	0.33	0.21	0.33	0.19
	TEMP-MED	0.36	0.28	0.32	0.19
	COMP-INI	0.34	0.26	0.46	0.29
	COMP-MED	0.29	0.22	0.32	0.23
RRT (Std.Err.) in ms	TEMP-INI	84 (10)	68 (11)	97 (12)	38 (7)
	TEMP-MED	88 (10)	71 (10)	87 (11)	65 (13)
	COMP-INI	101 (13)	81 (13)	132 (13)	113 (16)
	COMP-MED	86 (13)	57 (10)	100 (14)	79 (16)
TFT (Std.Err.) in ms	TEMP-INI	319 (14)	341 (20)	352 (18)	364 (20)
	TEMP-MED	317 (13)	332 (17)	340 (16)	388 (25)
	COMP-INI	353 (16)	342 (21)	376 (17)	486 (28)
	COMP-MED	328 (18)	300 (15)	367 (20)	413 (28)

comp-ini and comp-med differ more strongly than those of temp-ini and temp-med, suggesting an interaction between the fixed factors. Inferential statistics on these measures, however, do not yield any significant effect (all t-values are distinctly <|2|). Likewise, GLMMs on SKIP and RRP do not yield any significant effects (all z-values are distinctly <|2|, p>0.05). However, RegrP gives rise to a significant main effect for the factor 'disambiguation'. A regression was made from this word significantly more frequently when the disambiguating region required the temporal reading of *mehr* (*coeff. estimate=-0.2097, std.err.=0.1038, z-value=-2.02, p-value=0.043*). The main effect for 'verb stress' and the interaction term remain non-significant. The higher regression probability in the temp-conditions is reminiscent of the clause wrap-up effect that has been reported in Rayner et al. (2000). Its implications will be discussed in the discussion section together with the results from the other regions of interest.

Region 1: 1st word of disambiguating clause

The first word of the disambiguating clause (the actual disambiguating word) is a short function word in all conditions (3–4 characters). During first pass, it was skipped on average in 46% of the trials. Considering also later fixations, it was fixated at least once in 73% of the trials altogether. To test whether skipping of this word is affected by any of the controlled factors, a GLMM was fitted (with first pass skipping as the binomial response variable) yielding no significant effects for the factors 'disambiguation' (*z-value=1.486, p=0.137*) or 'verb stress' (*z-value=0.618, p=0.536*) or for the interaction term (*z-value=-0.803, p=0.422*). The evaluation of the reading times (FPRT, RRT and TFT) against the controlled factors plus the interaction does not yield any significant effect (all t-values <|2|). Similarly, RegrP and RRP lack significant effects (with z-values <|2| and p>0.05 for all main effects and interactions). Given the high skipping probability, the reading measures on this word may be unreliable.

Region 2: 2nd word of disambiguating clause

Because of the inconclusive and likely unreliable results on the disambiguating word, reading times on the second word of the disambiguating clause were examined, too (cf. Table 7 for the means). On average, the second word was skipped in 26% of the trials. In the comparative reading (comp-ini and comp-med), this word is a short function word (determiner, preposition or pronoun) in the majority of cases;[8] as for conditions temp-ini and temp-med, the word category of this position is more varied across items. Condition comp-med displays the highest skipping probability. Inferential statistics reveal that the interaction between 'disambiguation' and 'verb stress' approaches significance. Further analysis demonstrates that skipping occurred significantly more frequently in comp-med

[8]Two comparative items feature an adverb in this position.

as compared to COMP-INI (factor 'verb stress' (comparative conditions only): *coeff. estimate=-0.2532, std.err.=0.1136, z-value=-2.229, p=0.0258*). In contrast, the difference in skipping rate between TEMP-INI and TEMP-MED is negligible. The source of the interaction effect is therefore attributable to the difference between conditions COMP-INI and COMP-MED. While FPRT and RegrP do not show any considerable differences between the four conditions, the other measures reveal that the stress clash condition COMP-INI gives rise to the highest RRP, the highest RRT, and the highest TFT of the four conditions (cf. Table 7) in this region. No systematic effects of the fixed factors were found for TFT in this region (cf. Table 8). Inferential statistics for RRP show that the interaction between the factors 'verb stress' and 'disambiguation' approaches significance. Further analysis shows that re-reading this region is significantly more likely in the clash condition COMP-INI compared to condition COMP-MED (factor 'verb stress': *coeff. estimate=0.32, std.err.=0.10, z-value=3.11, p=0.0019*), whereas the difference between TEMP-INI and TEMP-MED is not systematic. This confirms again that the disparity between COMP-INI and COMP-MED is the main source of the interaction. As for RRT, the interaction between the fixed factors closely approximates significance. Singling out the two comparative conditions COMP-INI and COMP-MED, a linear mixed model confirms that the factor 'verb stress' significantly contributes to the difference between the two conditions (*coeff. estimate=0.799, std.err.=0.267, t-value=2.99*). Again the difference between TEMP-INI and TEMP-MED is marginal.

Region 3: Last word of disambiguating clause

The last word of the sentence was examined in order to determine whether the experimental factors show effects beyond the immediate vicinity of the disambiguating word. The average skipping probability is 19%. It is again the clash condition COMP-INI that displays the highest values in all other measures under scrutiny (cf. Table 7). While the values in COMP-INI and COMP-MED differ considerably, the values of TEMP-INI and TEMP-MED are much more similar. FPRT does not show any significant effects. Inferential statistics on regression probability reveal significant main effects for the factors 'disambiguation' and 'verb stress', as well as for the interaction (cf. Table 9). Closer inspection indicates that the two main effects are largely due to the salient values of COMP-INI. Looking specifically at the difference between the two COMP-conditions, a GLMM yields a significant effect for the factor 'verb stress' (*coeff. estimate=0.3386, std.err.=0.1065, z-value=3.179, p=0.0015*). As for TFT, the interaction between disambiguation and verb stress is also significant. Again, focusing on the COMP-conditions, the linear model yields a significant effect of 'verb stress' (*coeff. estimate=0.0917, std.err.=0.0313, t-value=2.931*). RRT and RRP give rise to a significant main effect for the factor 'disambiguation' with higher RRTs and RRPs for COMP-conditions compared to

Table 3.8: Modeling results for 2nd word of disambiguating clause (Region 2).

Measure	Coefficient	Estimate	Std. Error	Test Statistics
	disambig.	0.13893	0.09278	z=1.497, p=0.1343
SKIP	verb stress	-0.10734	0.08080	z=-1.328, p=0.1840
	disamb × v-stress	-0.14148	0.08031	z=-1.762, p=0.0782
	disambig.	-0.00084	0.01691	t=-0.05
FPRT	verb stress	-0.00803	0.01701	t=-0.47
	disamb × v-stress	-0.01308	0.01694	t=-0.77
	disambig.	-0.08538	0.09903	z=-0.862, p=0.389
RegrP	verb stress	-0.07302	0.09783	z=-0.746, p=0.455
	disamb × v-stress	0.04291	0.09744	z=0.440, p=0.660
	disambig.	0.17492	0.07782	z=2.248, p=0.0246
RRP	verb stress	0.17800	0.07378	z=2.413, p=0.0158
	disamb × v-stress	0.13826	0.07337	z=1.884, p=0.0595
	disambig.	0.20837	0.09024	t=2.309
RRT	verb stress	0.20739	0.08370	t=2.478
	disamb × v-stress	0.14613	0.08323	t=1.756
	disambig.	0.04257	0.02431	t=1.75
TFT	verb stress	0.01851	0.02105	t=0.88
	disamb × v-stress	0.00977	0.02099	t=0.47

Table 3.9: Modeling results for last word of disambiguating clause (Region 3).

Measure	Coefficient	Estimate	Std. Error	Test Statistics
SKIP	disambig.	-0.03024	0.10452	z=-0.289, p=0.772
	verb stress	-0.06954	0.09162	z=-0.759, p=0.448
	disamb × v-stress	-0.13224	0.09128	z=-1.449, p=0.147
FPRT	disambig.	0.00915	0.02845	t=0.32
	verb stress	0.02394	0.02338	t=1.02
	disamb × v-stress	0.01587	0.02319	t=0.68
RegrP	disambig.	0.25837	0.10327	z=2.502, p=0.0124
	verb stress	0.19447	0.08944	z=2.174, p=0.0297
	disamb × v-stress	0.14439	0.08850	z=1.632, p=0.1028
RRP	disambig.	0.31979	0.11060	z=2.891, p=0.00383
	verb stress	0.09508	0.09305	z=1.022, p=0.30686
	disamb × v-stress	0.10872	0.09251	z=1.175, p=0.23989
RRT	disambig.	0.26367	0.08293	t=3.180
	verb stress	0.06013	0.06712	t=0.896
	disamb × v-stress	0.12737	0.06664	t=1.911
TFT	disambig.	0.04404	0.02890	t=1.52
	verb stress	0.02828	0.02228	t=1.27
	disamb × v-stress	0.04875	0.02207	t=2.21

TEMP-INI and TEMP-MED.

In summary, the eye-movement data shows significantly increased reading costs for the comparative conditions COMP-INI and COMP-MED, when compared with the TEMP-conditions. Over and above this main effect of ‘disambiguation’, lower skipping probabilities, longer reading times and a higher likelihood of regressions are attested for the clash condition COMP-INI compared to the rhythmically innocuous condition COMP-MED, in the absence of a similar difference between the control conditions TEMP-INI and TEMP-MED. Although no significant effects were found on the actual disambiguating word (arguably due to the high skipping rate), the predicted interaction between the controlled factors ‘disambiguation’ and ‘verb stress’ is attested on the second word (in SKIP, RRT and RRP) and continues to affect eye-movements until the end of the sentence (RegrP, RRT and TFT). Note that, with the exception of FPRT and RegrP in Region 2, the coefficients of the predicted interaction between ‘disambiguation’ and ‘verb stress’ all signal higher reading costs for COMP-INI as compared to COMP-MED, i.e. they are negative for SKIP and positive for the other measures (cf. Tables 8 and 9). Even in the absence of significant effects for some of the dependent variables, this consistency suggests that the salience of the clash condition COMP-INI is systematic.

3.4.3 Discussion

As hypothesized, the present results suggest that rhythmic preferences indeed affect the silent parsing of written text. The eye-movement record in the disambiguating region of the test sentences attests systematic reading costs for the comparative disambiguation, reflecting the general preference for the temporal, unaccented version of *mehr*. The reading costs for comparative *mehr* are particularly high when the critical verb following *mehr* bears initial stress (COMP-INI). The prosodic representation of the comparative reading of *mehr* requires an accent on this word. Initial stress on the immediately following verb, therefore, would force a stress clash in this condition. The increased reading times relative to the rhythmically innocuous conditions indicate that the stress manipulation is critical for the assignment of syntactic structure. Readers avoid implicit accentuation of *mehr* when this would generate a stress clash. Accordingly, the unaccented temporal analysis of *mehr* is eminently preferred in this situation, which leads to increased processing demand if the comparative reading turns out to be the correct one.

The evaluation of several eye-tracking measures at different points within the disambiguating region supports this interpretation of the results. Before reviewing the supporting evidence, we address the inconclusive results that were obtained for the actual disambiguating word. This word is a short function word (3–4 characters) that introduces either a sentential complement to the preceding verb complex (in the temporal disambiguation) or a comparative complement starting with the word *als*. The shortness together with the fact that these words invariably introduce a new clause may be the reason for the high number of missing fixations on this word. Generally, short function words are heavily susceptible to skipping (Rayner, 1998). Moreover, it has been established that readers make relatively long saccades into a new clause, thus increasing the likelihood of skipping phrase initial words (Rayner et al., 2000). Together, these factors might well explain the missing fixations on the disambiguating word. It has to be noted though that fixating a word is not a necessary condition for processing it. Especially short words with a high frequency may be sufficiently recognized in parafoveal view. The reading data on the word preceding the disambiguating region (Region 0) provides a slight indication that the disambiguating word is already processed at this position: in line with the predictions, FPRT, RRT and TFT are (non-significantly) higher in the comparative conditions with the highest values in the clash condition COMP-INI. Regression probabilities in Region 0 are higher in the conditions with temporal reading of *mehr* as compared to the COMP-conditions. The regression probabilities for conditions TEMP-INI and TEMP-MED are strikingly reminiscent of the clause wrap-up effect reported by (Rayner et al., 2000, p. 1072); such an effect might be expected given the clause break at this position. However, why this wrap-up effect does not appear in conditions COMP-INI and COMP-MED remains open to speculation. One reason might be that readers experienced difficulties associated with the disambiguating word (Region 1) in parafoveal view in the COMP-conditions, which prevented them from programming a regressive saccade.

The second word of the disambiguating clause gives rise to an interaction between 'disambiguation' and 'verb stress' in skipping probability (SKIP), with significantly less frequent skipping in COMP-INI as compared to COMP-MED. As discussed above, effects concerning SKIP may reflect relatively early sentence processing stages. In fact, the decision to skip a word during first pass must be made while fixating a preceding region. It is therefore likely that readers make this decision while processing the actual disambiguating word.[9]

Also, re-reading probability and re-reading time are significantly increased in the clash condition COMP-INI. These measures are said to reflect more general comprehension difficulty (Clifton et al., 2007), suggesting that readers struggle to overcome the reading difficulties they encounter in this condition. It is also possible that these measures reflect reanalysis, which requires more effort in COMP-INI due to the stress clash. Similarly, on the last word of the disambiguating clause, the high regression probability and the high total fixation time indicate persisting reading difficulty in the clash condition COMP-INI.

The increased reading times for the comparative versions compared to the temporal disambiguation found in Regions 2 and 3 are most likely due to the general preference for the temporal reading (as would be predicted by Bader (1996a) and the unprepared oral reading experiment). However, since the lexical material between the two disambiguations is not necessarily comparable, this explanation should be taken with some caution.

In summary, the results of the silent reading experiment II appear to be compatible with the stress clash effect found in experiment I and thus confirm the involvement of supralexical, stress-based linguistic rhythm in parsing written text.

3.5 General Discussion

Previous research has uncovered effects of lexical stress on eye-movement patterns in silent reading. Those results were taken as an indication of an early speech-like prosodic representation of the text. The findings of the present experiments augment the evidence for a speech-like representation. They confirm that, during silent reading, readers mentally construct patterns of implicit lexical prominences

[9] Determining when exactly readers process the actual disambiguating word is not trivial given the high number of missing first pass fixations on this word and the likelihood that words are processed even in parafoveal view.

that evolve from the concatenation of individual words. While processing written text, readers obey prosodic-phonological preferences such as the principle of rhythmic alternation and they especially avoid sequences of adjacent stressed syllables. These results are important in that they demonstrate the involvement of supralexical, stress-based linguistic rhythm in silent reading of ordinary text. Moreover, the findings not only attest the mere existence of the rhythmic effect but also point to the functional role it may have in written sentence comprehension. The results shed light on the interplay of syntactic and prosodic processing during written sentence comprehension. The findings indicate that the process of analyzing the lexical-syntactic features of critical words is sensitive to the local prosodic environment even in the written modality, where no explicit prosodic cues exist. As for the timing of the rhythmic effect in relation to the syntactic analysis, two competing accounts will be discussed in the following:

First, it is conceivable that, once the reader encounters the ambiguous word *mehr*, he commits himself to the preferred temporal, unaccented reading and would only reconsider the decision if forced by syntactic counter-evidence in the disambiguating region. According to this view, the parser would initially disregard rhythmic preferences or prosodic cues in general (cf. Augurzky, 2006, for such a proposition). The increased reading costs for the conditions with the comparative disambiguation would come about solely due to the syntactic and concomitant prosodic reanalysis. Correspondingly, the prosodic reanalysis would force implicit accentuation of *mehr* and thus induce the rhythmically imperfect stress clash representation in condition COMP-INI, which would cause the additional increase in processing costs. While the effects in the 'late' reading measures (as, for instance re-reading time and total reading time) may be attributable to syntactic-prosodic reanalysis, the swiftness of the stress clash effect in silent reading (especially concerning skipping probability in Region 2) casts considerable doubt on the idea of the initial disregard of prosodic information. In order for the clash-effect to emerge only during reanalysis, the corresponding syntactic-prosodic revision would have to proceed very fast, i.e. probably immediately on encountering the disambiguating word. In contrast to this interpretation, syntactic-prosodic revisions are generally said to be rather costly and time-consuming (Bader, 1998). Also, this account cannot easily explain the results of experiment I. In the unprepared oral reading task, the choice of the accent status of *mehr* was clearly influenced by the immediate rhythmic environment, while, at this stage, disambiguating information further downstream was not systematically taken into account. Moreover, word prosodic information such as stress (the decisive factor for the evaluation of stress clash) is most likely computed rapidly online in silent reading (Ashby and Clifton, 2005; Ashby and Martin, 2008). Therefore, readers have the necessary information for evaluating potential stress clashes as soon as the critical words are combined within the reader's processing window, i.e. arguably before the evaluation of the disambiguating material in the present experiment.

The alternative account proposes that lexical stress contributes to parsing decisions more immediately. Hence, the eye-tracking effects may, but need not necessarily reflect prosodic reanalysis; instead, the early effects within the disambiguating region may be interpreted as reflecting integration difficulty. Accordingly, the interpretation of *mehr* would be affected already prior to the disambiguating region on the basis of the following verb's stress pattern. On this view, the parser is not always fully committed to the preferred reading when encountering the ambiguous word. Rather, the parser makes a variable choice (with a bias to the temporal reading), which depends on various sources of information (cf. van Gompel et al., 2001, and references therein). If – in spite of the general preference for the temporal, unaccented version – the parser happened to have initially foregrounded the comparative, accented analysis of *mehr*, it sets this analysis back when a stress clash would ensue due to the implicit accentuation of *mehr* and initial stress on the following verb.[10] Subsequently, the reader would expect a continuation compatible with the unaccented, temporal interpretation. The parser would therefore face integration difficulties when the disambiguating region forces the comparative interpretation of *mehr*, hence the increased processing costs in the clash condition COMP-INI. Significantly smaller processing costs ensue in the comparative condition with medial stress on the verb COMP-MED, as there is no rhythmic trigger that would demand unaccented *mehr* in addition to the general preference for the temporal analysis.

The very swiftness of the rhythmic effect in the eye tracking data makes this latter account more appealing. It is compatible with both the results from the unprepared oral reading experiment and the eye-tracking data (silent reading); moreover, it also acknowledges the finding that lexical-prosodic features are computed very early in reading (Ashby and Clifton, 2005; Ashby and Martin, 2008).

This interpretation of the results is consistent with the *Implicit Prosody Hypothesis* (IPH) by Fodor (2002), which predicts on-line effects of implicit prosody on the syntactic analysis. According to the IPH, the parser computes the most natural (default) prosody in line with the incremental analysis; the accruing prosody may affect syntactic decisions, biasing the parser to a syntactic analysis compatible with the prosodic representation. Previous experiments demonstrating on-line effects of implicit prosody on parsing (e.g. Hirose, 2003; Hwang and Steinhauer, 2011) were concerned with the interdependence of implicit prosodic phrasing and the attachment of larger syntactic constituents in reading. The prosodic trigger

[10] It is also conceivable that the verbal stress pattern already impinges on the initial processing of *mehr*, if it is assumed that the access to the syntactic features of *mehr* overlaps with access to prosodic-phonological features of the following verb. The reading data within the disambiguating region, however, are not informative on this matter.

in these studies, phrase length, requires at least low-level syntactic processing, namely the merging of words for the formation of phrases, whose length can be evaluated. In the present case, the default prosody (i.e. the preference for rhythmic alternation) was shown to affect more elementary building blocks of the syntactic structure, namely the assignment of the syntactic category of ambiguous words.

Beyond the resolution of the lexical ambiguity, the avoidance of accentuation of the critical word *mehr* in the face of a potential stress clash impinges – at least indirectly – on the syntactic predicate-argument structure. Note that the syntax-phonology interface in German requires arguments of verbs to be accented unless they are pronouns or given in the discourse (e.g. Truckenbrodt, 2006). Correspondingly, it is unlikely that readers interpreted *mehr* as a comparative filling the object position of the transitive verb when they did not accent it; conversely, accent on *mehr* might guide the parser to posit an argument.[11]

On the basis of these considerations, we suggest that prosodic-phonological and syntactic processing (and possibly semantic processing) are coupled and may alternately lead the way in written sentence comprehension. In the case of the temporarily ambiguous sentences of the present experiments, phonological well-formedness conditions like the principle of rhythmic alternation are considered for the computation of syntactic structure and may affect very early stages of the analysis – in this particular case: the determination of lexical-syntactic features.

This notion of written sentence comprehension is in line with models of sentence comprehension holding that different types of information (syntactic, semantic, phonological, etc.) may exert their influence on sentence comprehension as soon as they become available in the input (e.g. MacDonald et al., 1994; McRae et al., 1998; van Gompel et al., 2001). The present results demonstrate the influence of supralexical preferences concerning linguistic rhythm and accordingly call for the implementation of these effects in models of written sentence comprehension.

Finally, the results may also be taken as evidence for a more integrated account of sentence comprehension and production in reading. At the outset, it was predicted – based on speech production research – that (implicitly) producing a stress clash is generally avoided whenever more rhythmic alternatives are accessible. Both experiments confirm this view. Moreover, the results indicate that rhythmic preferences have repercussions for the comprehension process. Readers expect the written text, or more precisely the implicit phonological representation thereof, to be rhythmic; this apparently has consequences for the syntactic anal-

[11]Whether the stress clash directly affects syntactic computation, or whether it does so only via lexical disambiguation, cannot be determined on the basis of the present data. In any case, there is reason to assume that the domain of the stress clash effect is restricted to the phonological phrase (Hayes, 1989; Nespor and Vogel, 2007). Accordingly, stress clashes may possibly affect a syntactic constituent that corresponds to a phonological phrase (e.g. the constituent comprising a verb and its object) but not of larger syntactic clauses.

ysis. Thus, sentence comprehension in reading is at least in this respect driven by constraints that are standardly understood as being chiefly relevant to speech production.

We hope to have revealed the need to study the workings of linguistic rhythm beyond the acoustics of speech. Linguistic rhythm (and phonology in general) may be deemed an integral part of any linguistic behavior irrespective of the modality of perception and performance.

Chapter 4

Rhythmic Parsing

Abstract

The performance data from the experiment reported in Chapter 3 are modeled as an incremental constraint satisfaction process in the framework of an Optimality Theoretic parsing account. Solely making use of constraints derived from competence grammar, the model is capable of capturing the data and advocates the simultaneous application of syntactic, prosodic and syntax-phonology interface constraints in incremental processing. The model predicts that, in the case of syntactic indetermination, weak prosodic constraints may decide about syntactic ambiguity resolution. The performance-compatible OT grammar integrates the processes of syntactic parsing and prosodification in reading, hence dissolving the strict separation of language production and comprehension. At the same time the OT model endorses a bidirectional relationship between syntax and phonology in grammar.

4.1 Introduction

Readers generate from the graphemic string an intrinsic auditory version of the text entailing rich prosodic structure. Various reading studies have revealed that the silent prosodic rendition, called 'implicit prosody,' may affect the syntactic analysis of written text. Several studies on 'implicit prosody' suggest that readers' preferences concerning the prosodic representation compete with preferences with respect to the syntactic analysis. Prosodic preferences may be particularly forceful when there are no strong syntactic preferences. As an example, Fodor (1998) argues that the preference for roughly equal-sized prosodic phrases affects readers' attachment decisions in ambiguous environments like (1). In order to achieve a balanced output, a prosodic phrase boundary would separate *teacher's* and *mother* in (1-a) while, in (1-b), the genitive NP and the head noun are phrased together. Via syntax-prosody mapping constraints, this difference in prosodic phrasing impinges on the syntactic interpretation, with the adjective attaching to the genitive NP in (1-a) but attaching to the NP_{Gen}-NP complex in (1-b).

(1) a. (the cheerful teacher's) (mother-in-law)
    b. (the very cheerful) (teacher's mother)

However, the role of implicit prosody has been described as paradoxical (Fodor, 2002): on the one hand, it is assumed that prosody is shaped according to the syntactic structure assigned to the word string, suggesting that the syntactic analysis predetermines much of the prosodic representation; on the other hand, experimental evidence attests a clear influence of 'implicit prosody' on the syntactic analysis proper. Research on implicit prosody has therefore concentrated on the question concerning at what stage of the syntactic analysis phonological factors constrain the parse. Clearly, the balance principle responsible for the attachment preferences in (1) may only apply once all relevant words are in the parser's processing window. Correspondingly, Bader (1998) suggests a late influence of the prosodification on the syntactic analysis in reading. In his *Prosodic Constraint on Reanalysis*, it is proposed that prosodic factors add to the burden of syntactic reanalysis during sentence processing when the revised syntactic structure necessitates prosodic adjustments, too. Similarly, Augurzky (2006) claims that readers leave the prosodic rendition of the sentence underspecified during the initial processing stage, relying on purely syntactic cues. Only later are the syntactic parse and the prosodic rendition integrated. Hirose (2003) and Hwang and Steinhauer (2011) suggest that already during first pass parsing, syntactic analysis and prosodic representation are integrated, advocating early interaction of these domains in processing. Their experiments concern prosodic balance with respect to syntactic attachment preferences for long versus short phrases. It has to be noted, though, that the evaluation of phrase length by the parser requires the syntactic formation of these phrases in the first place. In this respect, the prosodification is dependent on at least limited syntactic pre-processing in these studies.

In the following section, we review experimental evidence that challenges the idea that readers build prosodic structure only on the basis of syntactic pre-processing. Conversely, the experiments suggest that, at points of syntactic underspecification, phonological constraints alone may guide syntactic structure assignment in reading. Beyond its psycholinguistic importance, such evidence has repercussions for the architecture of the competence grammar that the parser consults during processing. We argue that the grammar has to be devised in such a way as to allow phonological influence on syntactic structure assignment. On the basis of the empirical evidence, we propose and advocate a parsing model which makes explicit reference to an optimality theoretic competence grammar integrating constraints from the domains of syntax, phonology and the corresponding interface (section 3). We compare this competence-based model to other parsing models in section 4. The paper ends with a conclusion in section 5.

4.2 Experiment

Germanic languages have a general preference for the alternation of strong and weak syllables (e.g. Hayes, 1995; Liberman and Prince, 1977; Selkirk, 1984). It has been shown that a clash of two stressed syllables is avoided whenever more rhythmic alternatives are available. Faced with a potential stress clash, speakers might resort to stress shift (Bohn et al., 2011; Kiparsky, 1966; Visch, 1999) or they might choose a word order that prevents stress clash in the first place (Anttila et al., 2010; Shih et al., 2009). If, in silent reading, readers indeed generate a speech-like prosodic representation of the text, as proposed in Ashby and Martin (2008), clash avoidance should also hold in the written modality. That is, stress clash avoidance should have consequences for the syntactic processing of the written string: in the face of an ambiguous structure that involves a stress clash in one reading but not in the other, there should be a preference for the version without stress clash. This hypothesis has recently been confirmed in two reading experiments (Kentner, 2012, cf. chapter 3 of this dissertation), which we summarize briefly.[1]

The object of investigation was temporarily ambiguous structures, with two possible interpretations of the word *mehr* and specific prosodic representations for each of the syntactic interpretations (2). The word *mehr* is either the unaccented part of the temporal adverbial *nicht mehr* (Engl.: 'no longer')(TEMP) or the obligatorily accented, comparative quantifier (Engl.: 'more') (COMP). In order to

[1]For a detailed depiction of the results, the reader is referred to Kentner (2012) or chapter 3 of this dissertation.

test the rhythmic influence on syntactic parsing, the prosodically ambiguous word *mehr* was followed by a tri-syllabic verb with either initial (INI) or medial (MED) stress, yielding four experimental conditions. The sentences are disambiguated in the phrase that ends the sentence. In the TEMP-conditions, the disambiguating material is an extraposed sentential complement of the verb, introduced by a complementizer or wh-pronoun. The COMP-conditions are invariably disambiguated by the *als*-phrase, serving as the standard of comparison (i.e. the argument to the comparative *mehr*).

(2) Der Polizist sagte, dass man...
The policeman said that one ...

a. ... NICHT mehr NACHweisen kann, wer der Täter war. TEMP-INI
... couldn't prove anymore who the culprit was.

b. ... NICHT mehr erMITteln kann, wer der Täter war. TEMP-MED
... couldn't determine anymore who the culprit was.

c. ... nicht MEHR nachweisen kann, als die Tatzeit. COMP-INI
... couldn't prove more than the date of the crime.

d. ... nicht MEHR ermitteln kann, als die Tatzeit. COMP-MED
... couldn't determine more than the date of the crime.

In the case of the temporal adverbial *nicht mehr* (TEMP), the two graphemic words form a single lexical item since the meaning cannot be decomposed any further. Lexical stress (marked by single underlines) falls on *nicht*, and *mehr* remains unstressed. Correspondingly, the phrasal accent (marked by small caps) assigned to the adverb falls onto the stressed *nicht*. In adverb–verb sequences, the verb bears the main or nuclear accent (marked by small caps and double underlines) (Truckenbrodt, 2006). As for the COMP-conditions, *mehr* receives main phrase accent as it serves as a (comparative) complement to the verb; the verb itself may remain unaccented (Truckenbrodt, 2006).

Comparative, accented *mehr* followed by initial stress on the verb engenders a stress clash (COMP-INI). It was hypothesized that, without disambiguating information, readers should initially favor the unaccented, temporal interpretation of *mehr* in order to avoid the stress clash. Hence, in the clash-condition, readers should be forced to reanalyze when encountering the disambiguating region (the phrase at the end of the sentence) and thus experience increased processing demand.

First, an oral reading experiment (unprepared reading) was set up with 24 sets of sentences like (2). Twenty-four participants were asked to read single sentences from a screen without preparation. In order to prevent look-ahead to the disambiguating material at the end of the sentence, readers were asked to start reading out loud as soon as the sentence was displayed on screen. The accentuation patterns were evaluated by two student assistants, who were presented the sound files up to the verb. Their judgments confirmed the hypothesis that readers, as long as they were unaware of the disambiguating information, avoided accentuation of *mehr* when it was followed by a verb with initial stress. Specifically, readers accented 28% of *mehr* that were followed by a verb with medial stress but significantly fewer cases of *mehr* (19%) when a verb with initial stress followed. These numbers also reveal a strong general preference for the unaccented version of *mehr*.

Further analysis demonstrated that readers significantly slowed down in speech at the disambiguating region (i.e. in the phrase after the verb complex) when their accentuation of *mehr* turned out to be inappropriate relative to the disambiguating information. This slowdown is indicative of a garden path effect, suggesting that the readers' decision for accentuation of *mehr* involved a syntactic commitment to the relevant reading.

In a second experiment, 48 participants read the same set of sentences silently on screen with an eye-tracking device monitoring the fixation patterns. The evaluation of the eye-movement record in the disambiguating region attests significantly higher reading costs for the COMP-conditions. Arguably, the difficulties associated with the COMP-conditions reflect the general preference for the unaccented, temporal version of *mehr* that was observed in the oral reading experiment. On top of this main effect, the experiment yielded a significant interaction: reading times were significantly increased in the stress clash condition COMP-INI as compared to the other, rhythmically innocuous conditions (see Fig. 1).

The results are interpreted with recourse to the preference for rhythmic alternation, which apparently prevails even in silent reading. That is, the general preference for the unaccented, temporal-adverbial reading is reinforced by the rhythmic environment in the COMP-INI condition. In order to avoid a potential stress clash, readers avoid an 'implicit' accent on *mehr* and, correspondingly, follow the generally preferred temporal-adverbial reading. The temporal-adverbial reading, however, is incompatible with the disambiguating information, which causes the observed reading difficulties in the COMP-INI condition. Since, in the COMP-MED condition, there is no danger of stress clash, the preference for the temporal-adverbial reading is significantly weaker. The results support the claim that the

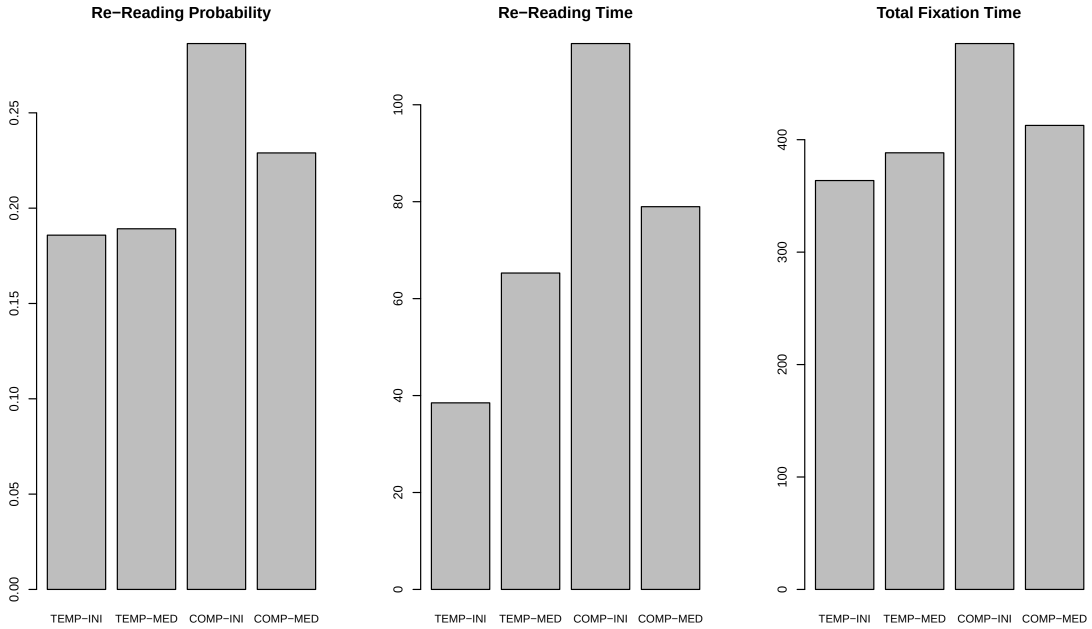


Figure 4.1: Re-reading probability, re-reading time and total fixation time within the disambiguating region broken down by condition. The highest reading times and fixation probabilities were obtained for the stress clash condition COMP-INI.

direct rhythmic environment, i.e. the lexical stress on the verb, affects the earliest stage of syntactic processing, viz. the determination of the syntactic category of the preceding ambiguous item *mehr*.

4.2.1 Discussion

The processing data on the *nicht mehr* ambiguity present a challenge for standard sentence processing models. The evidence suggests that prosodic planning, and more specifically the avoidance of stress clash, makes readers systematically leave a potentially accentable word (implicitly or explicitly) unaccented when adjacent syllables in neighboring words already require prosodic prominence. It was further shown that leaving the (syntactically and prosodically) ambiguous word unaccented has consequences for the parsing process. If disambiguating material later in the sentence requires an accent on the preceding ambiguous word, various measures of reading behavior point to processing difficulties, indicating that the syntactic analysis is directly conditioned by the prosodic rendition of the sentence in reading.

The prosodic effect on syntactic structure building seems to be immediate in the sense that it affects the earliest imaginable stage of syntactic analysis, namely the retrieval of the word's lexical-syntactic category. This state of affairs is incompatible with strictly feed-forward or unidirectional models of sentence processing and reading, in which prosodification is thought to occur only on the basis of the syntactic analysis (Kondo and Mazuka, 1996; Koriat et al., 2002; Wheeldon, 2000).

The findings also trigger questions about the architecture of the grammar that the processing mechanism consults. Since 'a performance model must certainly incorporate a grammar' (Chomsky, 1965, p.141), the grammar should be devised in such a way that it can be incorporated into a performance model. Correspondingly, a grammar that offers 'operational plausibility' (Lamb, 1998) should a priori be favored over one that draws hard boundaries between linguistic knowledge and its application in performance (Jackendoff, 2003; Sag and Wasow, 2011).

Particularly, the status of the syntax-phonology interface in grammar is at issue in the current case. Two competing accounts concerning the organization of this interface can be differentiated.

First, the standard (Chomskyan) view in generative linguistics assumes a unidirectional relationship between syntactic and phonological structure building, the former strictly controlling the latter. On this view, phonology merely interprets the conditions an autonomous syntactic module imposes on it (Chomsky, 1965, 1995). Phonological influences on the putatively 'central' syntactic component have been excluded on principled grounds (Pullum and Zwicky, 1988). This notion of unidirectionality parallels strictly feed forward language production models (e.g. Garrett, 1975). It is interesting to note that even prominent works in metrical stress theory (Hayes, 1995; Liberman and Prince, 1977; Kager, 1989; Selkirk, 1984), the subdiscipline studying the relationship between syntax and phonology and explicitly addressing the case of stress clash, do not question the assumed unidirectionality. It is difficult to conceptualize how a model that prohibits phonological influence on the syntactic component should deal with the processing data reported above.

The competing account on the syntax-phonology interface, proposing mutual interaction between these modules, seems more plausible. For example, Inkelas and Zec (1995), Zec and Inkelas (1990), Rice and Svenonius (1998) and Schlüter (2003) present a number of facts suggesting that phonological constraints may provoke the deviation from canonical constituent structure at least as long stronger syntactic forces remain indifferent to the resulting representation.[2] Most scholars, however, agree that the prosodic-phonological influence on syntax is more limited than syntactic influence on prosodic structure.

[2] In the exceptional case of poetry, otherwise dominating syntactic constraints may become subordinate to phonological forces (Golston and Riad, 2000), supporting the case for interaction of syntax and phonology in language.

In the following, I will outline a model of language competence that allows at least limited interaction of syntax and phonology and, at the same time, is capable of reproducing the principal results of the above reading experiments as an incremental parsing process. This model is, in effect, an optimality theoretic grammar (Prince and Smolensky, 1993, 2004) that is applied to sentence processing.

4.3 An Optimality Theoretic processing model

4.3.1 Foundations of Optimality Theory

Optimality theory (OT) (Prince and Smolensky, 1993, 2004) is a model of linguistic competence in which linguistic *input* structures are mapped onto linguistic *output* structures by means of two functions. The first function *GEN* generates a potentially infinite set of candidate outputs for a given input among which the grammatical one is chosen by the function *EVAL*. The structure of the candidates to be evaluated is hardly restricted. In fact, every candidate structure that is the result of any combination of linguistic primitives is taken to be an imaginable output. This assumption is referred to as the *richness of the base* (McCarthy, 2002). In order to select the grammatical output from this rich candidate base, the function *EVAL* applies. *EVAL* makes use of an arguably universal set of generally violable constraints *CON* in which simple and general principles are formulated. Certain constraints regulate the well-formedness of the output with respect to a given linguistic principle. As an example, structures that allow for rhythmic alternation of stressed and unstressed syllables are favored over rhythmically offending candidates. Other constraints control the conformity of input and output structure. Any deviance from the input specifications in the output candidates is punished by these *faithfulness* constraints. If in the input specifies a sequence of stressed syllables this should be represented in the output as well.

The linguistic principles formulated in the constraints thus impose conflicting requirements on the candidate outputs. No candidate can possibly satisfy all constraints. Instead, the candidate incurring the least serious violations of the constraint set is chosen as the 'winner' (Kager, 1999). The severity of the violation is evaluated by a conflict resolution device, i.e. the language specific constraint hierarchy in which the constraints are assigned different ranks. Violations of lower ranked constraints are tolerated as long as higher ranking constraints are satisfied. That way, the candidate that optimally satisfies the constraint hierarchy is taken as the grammatical output.

Connectionist foundations of OT

OT is a symbolic theory that is actually couched within a hybrid cognitive architecture that aims at integrating rule-governed symbolic processing with connectionist subsymbolic processing (Smolensky and Legendre, 2006). Subsymbolic processing in this architecture is implemented in terms of a neural network consisting of multiple interconnected units. The input to this network is understood as a pattern stimulating the spreading of activation through the network to produce a single output activation pattern. The connections between the units and their respective polarity (excitatory or inhibitory) and strength determine the output pattern in regulating the activation flow in specific ways. The output of the network is thus the one that conforms best to the demands of the network given the specifics of the input. An interesting property of such a model is that it does not impose excessive demands on attention or memory. The evaluation by constraint satisfaction is passive in nature – activation spreads through the network modulated by the strength and polarity of the connections – in contrast to a resource-limited active search process.

The functions of OT may be understood as metaphors of the features of this connectionist network. The connections and their specific weights represent the (conflicting) constraints and their respective ranks within the symbolic architecture. The imaginable output activation patterns of this network are subsymbolic descriptions of the symbolic candidate set.

The subsymbolic network thus does not produce an infinite number of outputs to choose from as a naïve interpretation of *GEN* in OT would suggest. Rather, the competing candidate outputs are merely *virtual* structures illustrating the network's conflict resolution behavior. They are not psychologically real as long as their specific activation pattern is actually realized by the network. Therefore, determining the optimal candidate does not involve the psychologically and computationally unrealistic notion of selecting one item in an infinite search space.

Although the functions of OT can be translated into subsymbolic neural networks, a qualitative difference between OT grammars and neural network implementations remains: In contrast to the fairly unrestricted possibilities concerning the weighting of connections in neural networks, common OT assumes strict constraint domination, that is 'no amount of success on a weak constraint can compensate for failure on a stronger one' (Prince and Smolensky, 1997). The reason for this regimentation lies in the typological observation that 'not every weighting of linguistic constraints produces a possible human language' (Prince and Smolensky, 1997). For the modeling of gradient data from psycholinguistic experiments, however, it might be necessary to revoke or at least relax this strong restriction. We will return to this issue in the discussion section.

4.3.2 Integrating comprehension and production in an Optimal Parsing account

Optimality Theory (OT) as a theory of grammar was originally designed to explain the relation between assumed underlying structures (INPUT) and surface structures (OUTPUT) in phonology. Input and output can be construed in various ways, however, depending on the nature of the task the grammar is used for.

Crucially, the same grammar can be applied to all kinds of linguistic performance. The grammar is independent of input and output modality but it is central to both. For example, in OT accounts that deal with language comprehension, the input is understood as the phonetic signal the listener is exposed to and the output is the mental representation he builds or accesses on the basis of the signal (e.g. Smolensky, 1996; Beaver and Lee, 2004). Similarly, OT can be applied to reading. In reading, the written string of lexical items can be construed as input and the output might be a syntactic description assigned to the string or the phonological representation thereof, which could be used for oral reading.

The OT grammar is thus a device that interprets (in the case of syntactic structure assignment) and generates (in the case of phonologization) linguistic objects; it is parser and synthesizer at the same time. Of course there is no perfect one to one relationship between the parser or synthesizer on the one hand and the grammar on the other. Parser and synthesizer are curtailed by factors that do not fall into the realm of grammar, such as situational context, time, memory, body and world knowledge. Core grammar abstracts away from the limitations these resources impose on the workings of the language faculty. The important influence of these factors on language performance is certainly acknowledged - and it might even be possible to formulate OT constraints that are attributed to them. However, the goal of the present model will be to bring together the parser and the grammar as close as possible, so the emphasis will be on the proper grammatical constraints.

The principle of constraint violability, which is central to OT, implies that the *degree* of adherence to the grammar can be evaluated for complete sentences as well as for partial, uncompleted, even ungrammatical candidate structures. This is a most desirable feature for any system of linguistic competence that aspires to have explanatory power for language performance. After all, being engaged in sentence comprehension or production means dealing with partial and imperfect linguistic objects at least as long as the sentence being processed is unfinished, arguably most of the time.

OT has the capability to build outputs for incremental input. It has been applied to sentence parsing and proven to have the capacity to model various facts from a wide range of sentence processing phenomena (Fanselow et al., 1999; Hoeks and Hendriks, 2011; Stevenson and Smolensky, 2006). The rationale is simple:

scanning the input incrementally from left to right (or in chronological order, in the case of auditory language comprehension), for any given piece of input, the OT parser would favor the analysis that best satisfies the constraint hierarchy.

Processing difficulty arises when high ranking constraints become relevant with additional input and force the parser to drop the currently optimal candidate in favor of a previously suboptimal one. The idea that constraints might become relevant with new input in parsing implicates that the order in which information is assessed by the parser is determined by the input stream. Depending on the structure of the input, constraints from different linguistic domains may interact freely at any time during the parse. That is, the evaluation process is cross-modular if the output is cross-modular. For example, constraints regulating the well-formedness of prosodic structure may interact with constraints formulating syntactic requirements in determining the structure of a linguistic object. In this respect, OT crucially differs from models of sentence processing that assume a fixed order in which different kinds of information are considered by the language processor at any given step of the analysis. For example, it is not the case in cross-modular OT that semantic and contextual information is only consulted after a first syntactic analysis of the input has been completed (as has been suggested in some models of sentence processing, e.g. Garden Path Theory, Frazier 1987, Friederici 2002) and a great deal of psycholinguistic evidence supports the parallel engagement of contextual semantic and syntactic knowledge in sentence parsing (e.g. Garnsey et al., 1997; Trueswell et al., 1994).

For the present case, the simultaneous application of constraints of various linguistic domains (syntax and phonology), the capability of evaluating incremental input, and the possibility to integrate both parsing and synthesizing of linguistic objects within a single framework are well suited for modeling the interaction of prosodification and syntactic structure building in reading performance. The OT model used for this endeavor has the following characteristics:

- The written string of lexical items serves as input.
- The output consists of both a syntactic description of the input and its prosodic rendition.
- The relevant constraints evaluate syntactic well-formedness, phonological well-formedness, and the agreement between syntactic and phonological structure.

The general method for testing whether OT can be used to model the parsing facts is as follows:

First, we determine the constraints relevant for the type of ambiguity under consideration. We make sure that the constraints are well motivated and have independent support. Second, we ascertain the ranking among the constraints as established in the relevant literature or on the basis of grammaticality judgments. Third, we verify the validity of the model by reference to the reading performance on the *nicht mehr*-ambiguity detailed in the previous section.

We assume that the *GEN* component of the model simultaneously generates two kinds of structured representations which are combined: a syntactic interpretation and a prosodic rendition. The written input provides information about the lexical items and their sequence. The combination of syntactic and prosodic structure is evaluated against the constraint hierarchy. The syntactic descriptions used here conform to standard assumptions concerning phrase structure, i.e. syntactic parses may generally be represented by a fully connected predicate argument structure with a single root representing the phrase or the sentence.

4.3.3 Motivating the constraints

In the case of the *nicht mehr*-ambiguity, the parser has the choice between the comparative reading of *mehr* and the adverbial reading of *nicht mehr*. In order to formulate the relevant constraints and their respective ranking, it is necessary to be clear about the syntax and prosody of the target structures to be evaluated.

The syntax of comparative *mehr* and adverbial *nicht mehr*

As is assumed for *more* in English, German comparative *mehr* is construed as a suppletive formation of an adjectival stem (*viel*, Engl.: *many*) and a degree marker (cf. Hendriks and De Hoop, 2001, for the relevant terminology), a.k.a. the comparative morpheme (*-er* in morphologically transparent forms).

Lechner (2003) notes that nominal comparatives like *mehr* in (3) display hybrid characteristics: they behave like DPs with respect to constituent distribution (as argument to the verb), but, at the same time, they have adjectival properties in providing the gradability necessary for comparatives. Obviously, the fused adjective-comparative morpheme selects a complex complement: As a prenominal adjective, it licenses the (optional) NP *Bier* (cf. Abney, 1987). In its function as degree marker, it selects the (obligatory) *als*-phrase containing the standard of comparison. Thus the relatively complex syntactic representation that follows Abney's and Lechner's analysis.

(3) Peter hat [$_{DP}$ [$_{DegP}$ [$_{AP}$ mehr [$_{NP}$ (Bier)]] als Susanne]] getrunken.
Peter has more (beer) than Susanne drunk
'Peter drank more (beer) than Susanne.'

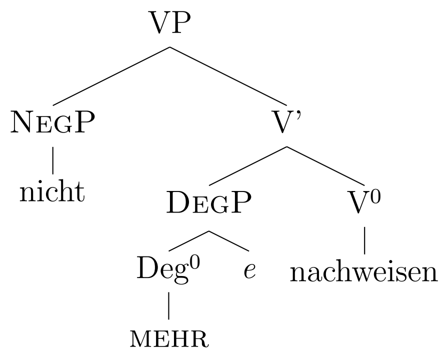
The fused adjective-comparative simultaneously serves as the head of the DegP, which is complemented by the *als*-phrase, and of the AP, which may take an NP as argument. Note that the NP within the AP does not necessarily need to be expressed in order for the sentence to be interpretable, cf. (3). On the other hand, the degree phrase introduced by *als* is required as the complement to the degree morpheme. Without a standard of comparison, the comparative cannot be interpreted, cf. (4). That is, leaving the standard of comparison unexpressed renders the sentence infelicitous if the context does not provide the information about the comparison.

(4) #Peter hat mehr getrunken.
Peter has more drunk.
'Peter drank more.'

In (5) the comparative *mehr* is the head of the DegP complement to the following transitive verb. Example (6) depicts the syntactic analysis of the region starting with *nicht* up to and including the verb. The negation is analyzed as an adjunct to the VP containing the DegP as complement to V^0. The DegP itself is elliptic as it lacks the degree phrase selected by the degree features of *mehr*, hence the empty sister node to Deg^0 (marked with *e*). The phrasal complement to the comparative head furnishing the standard of comparison in (5) is extraposed and surfaces only after the verb.

(5) Peter konnte nicht mehr nachweisen als die Blutwerte.
Peter could not more determine than the blood values
'Peter couldn't determine more than the blood values.'

(6)

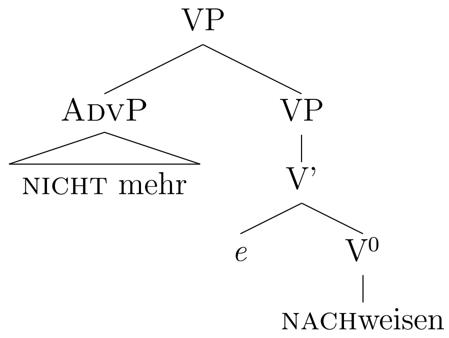


The adverbial counterpart of (5) has different syntactic properties. In this case (7), *mehr* is not gradable, and therefore cannot serve as the head of a degree phrase. Instead, *mehr* is part of a larger lexical unit, i.e. the adverb *nicht mehr*, which cannot be analyzed compositionally.

(7) Peter konnte nicht mehr nachweisen, dass die Werte erhöht waren.
Peter could not more determine, that the values increased were
'Peter couldn't determine anymore that the values were increased.'

As an adverb, *nicht mehr* adjoins to the VP. VP adjuncts, however, are not proper constituents of the core VP since, contrary to arguments, they are not selected by V^0. To reflect the fact that VP adjuncts are not exhaustively dominated by but are nevertheless part of the VP, an intermediate maximal projection is introduced in (8). This syntactic difference will be especially relevant for the prosodification of the phrase.

(8)



Observe that the representation in (8) lacks the argument that is required by the transitive V^0 (and hence marked by *e*). The argument is only realized postverbally as a sentential complement introduced by the complementizer *dass* in (7).

Comparing the two analyses in (6) and (8), they violate syntactic requirements at different points in the syntactic representation as long as they are unfinished. Focusing on the region up to and including the main verb, the comparative structure (6) lacks the complement of the degree head *mehr*, i.e. the degree phrase. It thus violates a syntactic constraint that demands that complements of heads, or arguments, surface. We will call this constraint Fill-Arg.[3]

(9) Fill-Arg: argument slots must be filled.

Note that, in the comparative reading, Fill-Arg is satisfied for the VP in providing the argument in the form of the DegP already pre-verbally. In the adverbial reading (7), again focusing on the material up to and including the VP, Fill-Arg is violated since the transitive verb lacks its argument. The argument to the verb appears only after the verb complex in the form of an extraposed sentential complement.

The violation of Fill-Arg is only relevant for the incomplete, unfinished representations in (6) and (8). In both structures, the postverbal material can be coindexed with the respective gap position (marked by *e*). That is, once the

[3]Applied to VPs, Fill-Arg has also been dubbed a violable version of the θ-criterion and accordingly been called ASSIGN-θ (Stevenson and Smolensky, 2006).

postverbal material is considered, the Fill-Arg violation will be suspended. However, it should be clear that the syntactic features of the postverbal material have to match the required features of the gap.

Prosodic properties of comparative *mehr* and adverbial *nicht mehr*

The comparative *mehr* is accented (as indicated by small caps in the tree diagram in (6)). Accentuation is regulated by the syntax-phonology interface constraint StressXP (Féry and Samek-Lodovici, 2006; Truckenbrodt, 1995, 2007; Truckenbrodt and Darcy, 2010). StressXP interprets the syntactic structure assigned to the input and requires that maximal projections (XPs) of lexical heads receive prosodic prominence on the level of the Phonological Phrase (PhP) by way of accentuation.

(10) StressXP: Each lexically headed XP contains a phrasal stress.

Though standardly understood as a functional projection, the DegP in (5) is headed by a degree operator *mehr* that has lexical content (due to its fusion with adjectival material). Therefore, StressXP applies and assigns accent to *mehr*. Note that with an accent on *mehr*, StressXP is satisfied for both the (lexical) DegP and the VP because the DegP is a proper constituent of the VP. Consequently, the verb may remain unaccented without violating StressXP. According to standard assumptions about X-bar structure, both VP adjuncts and VP arguments are considered maximal projections. However, adjuncts and arguments differ with respect to their affiliation to the VP: while VP arguments are proper constituents of the VP, adjuncts are not exhaustively dominated by the VP, as represented by the intermediate maximal projection in (8). StressXP is sensitive to this syntactic difference (see also Truckenbrodt, 2006, 2007). Prosodic prominence on a VP argument, in the present case on *mehr* in (5), implies prominence for the whole VP thus satisfying StressXP for the VP. Prosodic prominence on a VP adjunct, however, does not suffice to satisfy StressXP for the VP: the lower maximal projection, that is the core VP, also calls for an accent. Therefore, both adjunct and verb receive an accent in (8).[4] The adverb *nicht mehr* features lexical stress[5] on *nicht* so accent is correspondingly realized, leaving *mehr* without prosodic prominence in the temporal reading. StressXP is violated if the above conditions are not met, e.g. if adjunct or argument or the whole VP remain un-

[4]Matters are probably more complicated than described here: Féry and Herbst (2004) attest unaccented verbs in adjunct-verb sequences. The general tendency, however, is captured by the above account.

[5]As a temporal adverbial, *nicht mehr*, although it incorporates two graphemic words, is considered a single lexical entry as its meaning cannot be computed compositionally from its two constituents.

accented. This may happen under certain information structural conditions that will be briefly discussed in the following section.

In addition to StressXP, relative prominence of accents is regulated by Rightmost. This constraint simply requires that among the accents assigned to XPs the rightmost accent is strongest. Rightmost is equivalent to HeadPhrase (HP/HI) (Féry and Samek-Lodovici, 2006), which demands that heads of prosodic or intonational phrases appear close to the right edge. Accordingly, the accent on the adverbial *nicht mehr* in (8) is weaker than the accent on the verb (e.g. Féry and Samek-Lodovici, 2006; Truckenbrodt, 2007).

(11) Rightmost: The rightmost accent within a prosodic or intonational phrase is strongest.

Three more constraints are relevant for the present case. These impose restrictions on the prosodic-phonological representation independent of its syntactic structure. The requirements of StressXP are in conflict with a purely phonological constraint, *Accent. This constraint penalizes prevalent accentuation by assigning each accent a violation mark. It is motivated by the fact that a sentence like (12) is not normally uttered with accent on each word. An accentuation pattern of this type may only be licensed under very strong pragmatic conditions.

(12) L*H L*H L*H H*L%
Peter wollte die Kosten ermitteln.
'Peter wanted to calculate the costs.'

In addition, two further constraints play a role in deriving the prosodic structure of phrases and sentences. *Clash (Prince, 1983; Selkirk, 1984) militates against adjacent stressed syllables within a phonological phrase. The workings of *Clash are exemplified in (13). The accented monosyllabic *Hemd* is adjacent to *anziehen*, which bears main stress on the initial syllable in its citation form (13-a) (relative prominence is indicated by underscoring). According to Bohn et al. (2011), Giegerich (1985), Kiparsky (1966), and Visch (1999), the stress position of the trisyllabic main verb may shift away from the first syllable to the second in such a situation, yielding the stress pattern in (13-b). Stress shift warrants avoidance of stress clash, satisfying *Clash. However, this procedure violates another constraint that requires faithfulness to the citation form stress pattern, which, in the case of *anziehen*, would demand initial stress. The latter constraint will be called IdentStress.

(13) a. #... das Hemd <u>an</u>ziehen.
b. ... das Hemd an<u>zie</u>hen.
... the shirt on-take

'... to put on the shirt.'

Note that stress clashes cannot always be resolved. In (14), there is simply no other syllable to which stress could shift, so (14) will always incur a violation of *CLASH.

(14) ...weil er das HEMD holt.
...because he the shirt pick-up.
'...because he is picking up the shirt.'

Taken together, the constraints relevant for the present case are of three types: FILL-ARG regulates syntactic structure building. STRESSXP supervises the syntax-phonology interface. Finally, four further constraints are responsible for the prosodic-phonological representation, namely RIGHTMOST, *ACCENT, *CLASH, and IDENTSTRESS.

4.3.4 Determining the constraint hierarchy

In the previous section, the relevant constraints and their requirements were introduced. This section motivates the constraint hierarchy which ultimately governs the parsing process. It is customary in most applications of Optimality Theory to determine the ranking of constraints on the basis of the empirical facts alone. In our case, the empirical facts correspond to the parsing preferences as evidenced by the reading experiments on the *nicht mehr*-ambiguity. However, we need independent motivation for the constraint ranking since we are not only interested in showing that OT can be used to model the parsing and prosodification facts in reading, but also that the processing facts can be modeled with the same grammar that describes and explains the structure of linguistic objects in general. We therefore consult the competence grammar to ascertain the constraint hierarchy and test this ranking against the performance data from the reading experiment.

The syntactic constraint FILL-ARG is, in essence, one side of the θ-criterion, demanding that argument slots or θ-roles be filled. For a sentence to be grammatical, the θ-criterion has to be fulfilled (Chomsky, 1981). Correspondingly, this constraint is undominated.

We turn now to the constraints regulating the syntax-prosody interface and the phonological representation.

STRESSXP may be violated under certain conditions without automatically inducing ungrammaticality. As an example, a VP may lack accentuation if it is already given in the discourse. This is the case in (15-b), where the locative adjunct *unter der Dusche* receives accent but the VP proper may remain unaccented due to discourse givenness. A pattern of this type may emerge as a result of high ranking constraints that guide the interface between information structure and prosody such as STRESSFOCUS and DESTRESSGIVEN (Féry and Samek-Lodovici, 2006).

(15) Where did Julie sing?

a. #Julie hat unter der DUSCHE GESUNGEN.

b. Julie hat unter der DUSCHE gesungen.
Julie has under the shower sung
'Julie sang under the shower.'

A violation of RIGHTMOST is tolerated under similar circumstances, i.e. if information structural requirements govern the prosodification. Consider the elliptical coordination structure in (16). Owing to the contrast, the verb *loves* in (16) is assigned strongest accentuation. The object to the right may be accented, too, but its accent is certainly weaker. Generally, post-nuclear accents violate RIGHTMOST.

(16) Mary DESPISES and Peter LOVES fruity CREAM tarts.

The OT literature on sentence intonation has established that STRESSXP dominates RIGHTMOST in German and English (Féry and Samek-Lodovici, 2006). We will stick to this independently motivated hierarchy.

The constraint *CLASH will be violated whenever the adjacency of two prominent or strong syllables cannot be resolved. Adjacency of strong syllables is a widespread phenomenon, as may be exemplified by the very conventionality of phrases like (17) consisting of three consecutive stressed syllables.

(17) ... weil Linn BROT kauft.
... because Linn bread buys
'...because Linn is buying bread.'

This suggests a rather low rank for *CLASH. IDENTSTRESS must be ranked even lower, otherwise stress shift in the face of a potential stress clash, as in (13) above, would remain inexplicable. *ACCENT is necessarily violated by every utterance if we subscribe to the assumption of the prosodic hierarchy according to which every utterance entails a phonological phrase which, by definition, is headed by an accent. *ACCENT will therefore hold the lowest rank among the relevant constraints.

The above considerations establish the ranking in (18).

(18) FILL-ARG ≫ STRESSXP ≫ RIGHTMOST ≫ *CLASH
≫ IDENTSTRESS ≫ *ACCENT

This ranking will be used throughout in the following demonstration of the model.

4.3.5 Putting the model to work

The constraint hierarchy will be applied to the *nicht mehr*-ambiguity introduced above. The OT-parser/synthesizer, going from left to right, takes as input the written lexical items and Gen incrementally assigns the string a set of syntactic and prosodic candidate descriptions.[6] For each parsing step, Eval choses among the candidate outputs the one that best harmonizes with the constraint hierarchy. According to OT convention, the optimal candidate will be marked with a pointing finger. For the application of OT to incremental processing, this convention will be expanded in the following way:

Incremental candidates that are optimal throughout the parsing process collect pointing fingers for each parsing step. That is, if the optimal candidate structure of parsing step *n-1* is consistent with the optimal candidate at parsing step *n*, that candidate will be marked with two pointing fingers. Conversely, once previously optimal candidates become suboptimal, the corresponding candidate is flagged with a hash mark. Candidates marked with a hash mark may still take part in the competition and new input may revive them as optimal candidates.

There are two ways to predict processing costs on the basis of this OT formalism. First, a parse incurs processing costs at parsing step *n* if the optimal candidate Φ is inconsistent with the optimal candidate at *n-1*; the costs increase with every previous parsing step *n-1...n-k* for which a partial structure consistent with Φ was suboptimal.

Secondly, and more importantly, the higher the rank of the constraint responsible for the failure of a previously optimal candidate, the higher the processing costs associated with it (Stevenson and Smolensky, 2006).

In what follows, we depict an incremental OT processing model that makes falsifiable predictions for reading performance. Crucially, the OT model integrates syntactic parsing (comprehension-oriented optimization) and prosodification (production-oriented optimization) using a single constraint hierarchy. The constraints of the model have their basis in grammar, so the reason for the preference of one output over the other is a grammatical one. Also, the interplay of production and comprehension is thus rooted in the OT grammar.[7]

For illustration, the model will be applied to the four conditions of the *nicht mehr*-ambiguity in (19).

(19) Peter konnte...
'Peter could...'

[6]In some models that use OT in language comprehension, the function generating the candidates is called Int for 'interpretation' (Stevenson and Smolensky, 2006). We keep using Gen and understand it as a cover term for both production oriented as well as comprehension oriented candidate generation.

[7]This is not to deny that performance constraints have a role to play in parsing; however, the purely grammatical model is more economical and thus yields a more comprehensive explanation of the processing results.

a. ...nicht mehr ausrechnen, dass... TEMP-INI
...not more calculate that...
'...not calculate anymore that...'

b. ...nicht mehr berechnen, dass... TEMP-MED
...not more calculate than...
'...not calculate anymore that...'

c. ...nicht mehr ausrechnen, als... COMP-INI
...not more calculate than...
'...not calculate more than...'

d. ...nicht mehr berechnen, als... COMP-MED
...not more calculate than...
'...not calculate more than...'

Three parsing steps will be considered, starting in the ambiguous region, namely with the two graphemic words *nicht mehr*. As shown in Tableau 1, the adverbial reading (candidates a–d) is preferred as it satisfies the high ranking FILL-ARG in contrast to the competing comparative reading. The latter reading is suboptimal as it lacks the complement to the comparative head *mehr*, thus incurring a violation of FILL-ARG (candidates e–h).

As for the prosodification there are four logical possibilities concerning the distribution of accents: accent on either *nicht* or *mehr* or accent on both or accent on neither (accents are marked by small caps).[8]

STRESSXP requires accentuation of the adverbial phrase. Accordingly, the accentless candidate (d) is out of bounds. Double accentuation in (c) and (f) is prohibited by *CLASH and also incurs a gratuitous violation of *ACCENT. Candidate (b) features accent on *mehr*, violating the lexical-phonological prominence pattern of the adverbial *nicht mehr*, which bears stress on the initial syllable. Candidate (b) thus founders on IDENTSTRESS. The optimal candidate (a) only incurs a single violation of *ACCENT.

The suboptimal comparative candidates all founder on FILL-ARG for lack of the comparative phrase (missing argument marked by $e_?$), but it is still possible to determine the best among the bad ones: candidate (e), featuring a single accent on *mehr*, fares best. Leaving comparative *mehr* unaccented incurs a violation of STRESSXP (g and h); double accentuation violates *CLASH and *ACCENT (f).

[8]For ease of exposition, the relative strength of accentuation is not evaluated at this parsing step, i.e. RIGHTMOST is ignored in Tableau 1. We simply assume that, in the case of accent on both *nicht* and *mehr* (candidates c and f), RIGHTMOST requires nuclear accent to fall on *mehr*. Note that doubly accented candidates are also suboptimal due to their violating *CLASH and *ACCENT.

/nicht mehr/		Fill-Arg	StressXP	Rightmost	*Clash	IdentStress	*Accent
a. ☞	[$_{VP}$ [$_{AdvP}$ NICHT mehr]						*
b.	[$_{VP}$ [$_{AdvP}$ nicht MEHR]					*!	*
c.	[$_{VP}$ [$_{AdvP}$ NICHT MEHR]				*!	*	**
d.	[$_{VP}$ [$_{AdvP}$ nicht mehr]		*!			*	
e.	[$_{VP}$ nicht [$_{DegP}$ MEHR $e_?$]	*!					*
f.	[$_{VP}$ NICHT [$_{DegP}$ MEHR $e_?$]	*!			*		**
g.	[$_{VP}$ NICHT [$_{DegP}$ mehr $e_?$]	*!	*				*
h.	[$_{VP}$ nicht [$_{DegP}$ mehr $e_?$]	*!	*				

Table 4.1: 1st parsing step

Turning to the second parsing step (Tableau 2), the OT processor considers *nicht mehr* together with the following verb. Given that the verb is obligatorily transitive, it incurs a violation of Fill-Arg in those candidate structures that do not provide a direct object. This holds for the adverbial reading since the adverb cannot occupy the argument slot. The nominal comparative *mehr*, however, may step in as argument to the verb. Still, the comparative parse continues to violate Fill-Arg as it lacks the complement of the degree head. Therefore, the adverbial and comparative readings both fail to satisfy Fill-Arg albeit for different reasons. As a result, both syntactic analyses are on equal footing concerning the high ranking syntactic constraints. The competition is therefore decided by lower ranking phonological constraints. In the case of a verb with lexical stress on the initial syllable (upper half of Tableau 2), the adverbial reading wins. The comparative reading with accented *mehr* and an initially stressed verb founders on *Clash (candidate e). The stress shift candidate (e') fails due to a violation of IdentStress.

As for the cases with the verb featuring medial stress (cf. lower panel of Tableau 2), the situation changes. In contrast to the corresponding candidate in the upper half of Tableau 2, candidate (e) with a stressless initial syllable does not incur a violation of *Clash because of the favorable stress pattern on the verb. Instead, the adverbial candidate with accent on both *nicht* and the verb (candidate a) incurs one more violation of *Accent than the optimal comparative candidate (e), which turns out to be fatal. The predecessor of the adverbial candidate (a) was the optimal candidate of the previous parsing step and is correspondingly flagged with the # mark. This change of parsing preferences from step 1 to step 2 hinges on a very low ranking constraint. It may therefore be all but a weak

		/nicht mehr ausrechnen/	Fill-Arg	Rmost	*Clash	IdStress	*Acc
a.	☞ ☞	[$_{VP}$ [$_{AdvP}$ NICHT mehr] [$_{VP}$ $e_?$ AUSrechnen]]	*				**
a'.		[$_{VP}$ [$_{AdvP}$ NICHT mehr] [$_{VP}$ $e_?$ AUSrechnen]]	*	*!			*
e.		[$_{VP}$ nicht [$_{DegP}$ MEHR $e_?$] ausrechnen]	*		*!		*
e'.		[$_{VP}$ nicht [$_{DegP}$ MEHR $e_?$] ausrechnen]	*			*!	*
		/nicht mehr berechnen/					
a.	☞#	[$_{VP}$ [$_{AdvP}$ NICHT mehr] [$_{VP}$ $e_?$ beRECHnen]]	*				*!*
a'.		[$_{VP}$ [$_{AdvP}$ NICHT mehr] [$_{VP}$ $e_?$ beRECHnen]]	*	*!			*
e.	☞	[$_{VP}$ nicht [$_{DegP}$ MEHR $e_?$] berechnen]	*				*
e'.		[$_{VP}$ nicht [$_{DegP}$ MEHR $e_?$] berechnen]	*		*!	*	*

Table 4.2: 2nd parsing step (candidates violating StressXP are ignored).

preference. Accordingly, processing costs associated with the preference change will be relatively low.

In the third parsing step (cf. Tableaux 3 and 4), the parser encounters the disambiguating material, i.e. the complementizer phrase introduced by *dass* or the comparative phrase with *als*. As discussed above, the complementizer phrase may fill the up to now open VP argument by providing a sentential complement to the transitive verb in the adverbial reading; at the same time, the complementizer is incompatible with the comparative version. Conversely, the *als*-phrase may be interpreted as the argument to the comparative DegP but it is impermissible in the context of the adverbial reading, which demands a complement to the VP. The outcome of the competition between the syntactic analyses therefore depends on the high ranking constraint Fill-Arg. To maintain clarity, we will only be considering the best candidates of each reading for the two parses.

In the case of the initially stressed verb followed by the complementizer phrase (Tableau 3, upper panel), the parser may simply maintain the analysis established in both previous parsing steps; the complementizer phrase introducing the VP complement can simply be coindexed with the VP argument slot. Fill-Arg is thus satisfied and no processing difficulty is predicted in this case. The best of the significantly worse candidates featuring the comparative reading was suboptimal throughout the parse and founders on Fill-Arg. The CP cannot legally fill the still open argument slot within the DegP, as *mehr* obligatorily selects the comparative phrase introduced by *als*.

Encountering the complementizer after a verb featuring medial stress (Tableau 3, lower panel), however, leads to another change of parsing preferences. Again, the high ranking syntactic constraints decide about the winning candidate. Observe that in the previous parsing step, the comparative reading was preferred over the adverbial analysis, albeit weakly. The very weakness of the parsing preference in step 2, together with the brevity of the period in which this preference has held, may alleviate the processing costs that the repeated preference change would predict for parsing step 2. Assuming a ranked parallelism in OT parsing, the temporarily suboptimal adverbial reading might still be active to a relatively high degree.

		TEMP-INI: /nicht mehr ausrechnen... dass/	FILL-ARG	IDSTRESS	*ACC
a.	☞ ☞ ☞	[$_{VP}$ [$_{AdvP}$ NICHT mehr] [$_{VP}$ e_i AUSrechnen]] [$_{CPi}$ dass			**
e'.		[$_{VP}$ nicht [$_{DP}$ MEHR $e_?$] ausrechnen] [$_{CP}$ dass	*!	*	*
		TEMP-MED: /nicht mehr berechnen... dass/			
a.	☞# ☞	[$_{VP}$ [$_{AdvP}$ NICHT mehr] [$_{VP}$ e_i beRECHnen]] [$_{CPi}$ dass			**
e.	☞ #	[$_{VP}$ nicht [$_{DP}$ MEHR $e_?$] berechnen] [$_{CP}$ dass	*!		*

Table 4.3: 3rd parsing step (disambiguation towards adverbial reading). Candidates violating STRESSXP, RIGHTMOST and *CLASH are ignored.

		COMP-INI: /nicht mehr ausrechnen... als/	FILL-ARG	IDSTRESS	*ACC
a.	☞ ☞ #	[$_{VP}$ [$_{AdvP}$ NICHT mehr] [$_{VP}$ $e_?$ AUSrechnen]] [$_{ComP}$ als	*!		**
e'.	☞	[$_{VP}$ nicht [$_{DP}$ MEHR e_i] ausrechnen] [$_{ComPi}$ als		*	*
		COMP-MED: /nicht mehr berechnen... als/			
a.	☞# #	[$_{VP}$ [$_{AdvP}$ NICHT mehr] [$_{VP}$ $e_?$ beRECHnen]] [$_{ComP}$ als	*!		**
e.	☞ ☞	[$_{VP}$ nicht [$_{DP}$ MEHR $e_?$] berechnen] [$_{ComPi}$ als			*

Table 4.4: 3rd parsing step (disambiguation towards comparative reading). Candidates violating STRESSXP, RIGHTMOST and *CLASH are ignored.

In the face of a comparative *als*-phrase preceded by the initially stressed verb (upper panel of Tableau 4), the parser is forced to revoke the previously preferred analysis (candidate a) due to the requirements of FILL-ARG. The processing costs associated with this change of preference towards candidate (e') should be high for two reasons: First, the now discarded adverbial analysis was established over both preceding parsing steps, suggesting relative stability of this parse. Second, the constraint responsible for the failure of the previously optimal candidate is a high ranking one. Note that this parsing step involves a stress shift on the verb that is required by *CLASH.

The OT parser predicts decidedly less difficulty in the case of the medially stressed verb (lower panel of Tableau 4). In the previous parsing step, a weak preference for the comparative reading has already been established. The new input simply confirms this preference for candidate (e).

4.4 Discussion

The performance differences the model predicts for the four conditions at parsing step 3 do indeed appear to reflect the reading data obtained in the experiments. Little processing difficulty is predicted for both conditions with the adverbial reading (Tableau 3). This is obvious for the condition with the initially stressed verb. The optimal candidate is syntactically and prosodically consistent with the optimal candidates at both previous parsing steps (hence the three pointing fingers). In the case of the verb with medial stress, the deviating parsing preference that is predicted for the second parsing step is a very weak one. It hinges on a low ranking constraint and might thus be easily overwritten in the third step.

Turning to the two conditions featuring the comparative reading (Tableau 4), a clear difference in processing difficulty is predicted between cases with initial and medial stress on the verb. In the latter condition, a weak preference that had been established at the second parsing step is confirmed. The structure should therefore be processed relatively easily compared to the comparative *mehr* followed by a verb with initial stress. It is this condition that is predicted to produce the highest processing costs: the competitor that was the optimal candidate in the two previous parsing steps founders on a high ranking constraint.

This prediction of the model is borne out in the actual data of the silent reading experiment. In the disambiguating region, several dependent variables reveal an interaction between the factors 'syntactic reading' and 'verbal stress pattern' that is mainly due to the striking reading difficulty observed for the condition that forces a stress clash (or else a deviation from the lexically determined stress pattern), i.e. comparative *mehr* followed by initial stress on the verb.

The model is also compatible with the general prevalence of unaccented readings of *mehr* that was observed in the oral reading experiment (only 25% of cases accented; cf. summary of oral reading experiment). The avoidance of accent on *mehr* can be explained with recourse to the model's first parsing step (cf. Tableau 1), in which the adverbial reading was established as the optimal interpretation, and the competing comparative version ruled out by strong syntactic constraints. Since, in the oral reading experiment, readers were encouraged to read as fluently as possible without self-corrections, this initial analysis may have been only mildly affected by the incoming verb in the second parsing step.[9]

Generally, the proposed OT parser accounts for the prosodic effects that were shown to act upon syntactic ambiguity resolution in reading. The OT model makes direct use of the grammar in determining its parsing preferences. The assumption is that grammatical restrictions are applied to incremental structure building and that preferences at points of ambiguity reflect these grammatical requirements. The constraint ranking that derives the parsing preferences has a sound and independent grammatical motivation and, in the present case, it can do without additional extra-grammatical processing constraints. Moreover, not only the constraints but also the specific ranking employed for this performance theory hold for the ordinary competence grammar. Importantly, this model is cross-modular in that it integrates constraints from different modules of grammar, namely syntax, phonology and the corresponding interface.

With that said, we shall briefly return to the problem of the syntax-phonology interface in grammar touched upon in the introduction. The present model is a performance-compatible grammatical device that allows phonological constraints to act upon syntactic structure building. This is especially evident in parsing step 2 (cf. Tableau 2, lower panel), in which the optimal candidate, complete with a syntactic description, is selected as optimal because it fares better than the relevant syntactic competitor only with respect to phonological constraints. Given that it is the OT grammar proper that produces this result, this parsing step constitutes an offense against the notion of phonology-free syntax as forcefully advocated by Pullum and Zwicky (1988). If this incremental model and its constraint hierarchy is realistic, the idea of a merely unidirectional relation between syntax and phonology is once more invalidated (see Inkelas and Zec, 1995; Rice and Svenonius, 1998; Schlüter, 2003; Zec and Inkelas, 1990, for more evidence against phonology-immune constituent structure building). Rather, this phenomenon is reminiscent of the idea that Bierwisch (1966) put forward referring to the role of prosody in language production:

> *Es ist sehr wohl möglich, dass ein Sprecher sich durch ein begonnenes Into-*

[9] This explanation – which is based on the incremental application of grammatical constraints – is orthogonal to the one offered by Bader (1998) in the context of a similar syntactic/prosodic ambiguity. He states that those readings are difficult to parse in which function words have to be assigned accent. That is, in contrast to the grammatical account defended here, Bader (1998) ascribes the attested preference pattern to the lexicon when stating that function words should preferably remain unaccented. As noted above, comparative *mehr*, even though it may project a functional XP, is not an indisputable function word as it clearly features adjectival properties.

> *nationsmuster zur Wahl einer bestimmten syntaktischen Struktur veranlasst fühlt* [...].
> It is very well possible that a speaker feels himself prompted to choose a certain syntactic structure owing to the intonation pattern that he has started.
>
> Bierwisch (1966), p. 105.; translation by the author

A grammar and/or processing model that does justice to phonological interference of this type must surely allow for at least limited interaction of syntax and prosody. Without further justification, I submit that situations of the kind found here (and formulated by Bierwisch, 1966) are not uncommon in incremental sentence processing. For the assignment of a syntactic description to a full sentence, however, recourse to phonology might be the exception rather than the rule.

In the following section I will compare the OT approach to other models of sentence processing.

4.4.1 Comparison to other sentence processing models

Other constraint-based approaches

The proposed OT processor has a lot in common with the family of constraint satisfaction models that have been formulated for sentence comprehension (MacDonald et al. (1994); Trueswell et al. (1994) and the computer implemented variants such as McRae et al. (1998); Tabor et al. (1997)). Just as the subsymbolic variant of OT (cf. Smolensky and Legendre, 2006), constraint satisfaction models of information processing are based on the idea of spreading activation in neural networks. Together with these models, OT assumes that multiple constraints from various domains may interact during ambiguity resolution.

As opposed to the present model, the standard constraint satisfaction models generally de-emphasize grammatical influences on parsing (cf. Stevenson and Smolensky, 2006), focusing on extra-grammatical factors such as lexical or structural frequency. Models like the Competition-Integration Model (McRae et al., 1998) do not provide an explicit testable competence grammar but rely on unspecified modules providing the syntactic alternatives to be evaluated. It is questionable what the cognitive equivalent of such modules is and what their relation to competence grammar might be. In short, the relationship between competence grammar and the processing component is underspecified in these models.

The architectures of the standard constraint satisfaction models seem to be task specific. In the psycholinguistic arena, distinct connectionist models have been proposed for speech processing (e.g. McClelland and Elman, 1986), sentence comprehension (Tabor et al., 1997; Tabor and Tanenhaus, 1999; McRae et al., 1998), reading aloud (on the word level only: Seidenberg and Plaut, 1998; Rastle and Coltheart, 1998), and language production (e.g. Dell et al., 1999; Bock and Griffin, 2000). As Christiansen and Chater (2001) point out, it is far from clear whether and how these models will eventually be integrated to cover full-scale human language performance. OT as a theory of grammar 'at least has the explicit aim of developing a general set of constraints that applies to all sentences in a language' (Archangeli, 1997, as cited by Hoeks and Hendriks (2011)); this general set of constraints and its ranking should be central to all kinds of linguistic tasks.

Owing to the strict constraint domination approach taken here, the present model does not generate numerical predictions of processing difficulty that could be evaluated against numerical processing data. The lack of this ability is clearly detrimental to the OT model in comparison to other constraint-based approaches. The relation of the model's prediction to the experimental data is discussed in more detail below (cf. 4.4.2).

Two-stage parsing accounts

Together with other constraint satisfaction models, OT assumes that multiple constraints from various domains may interact during ambiguity resolution. The kind of information assessed in constraint satisfaction models depends on the nature of the input alone. Whichever constraint is responsive to a given piece of input will be active in ambiguity resolution – independent of its grammatical domain. That is, constraint-based approaches assume parallel evaluation of several structural descriptions of the input.

This behavior contrasts with so-called *syntax-first* serial accounts like the Garden Path model in which syntactic structure is given temporal priority in evaluation (Frazier, 1987; Friederici, 1995, 2002). According to these models, comprehending a sentence involves several (at least two) stages, each of which is dedicated to the processing of a certain kind of information. Due to limited memory, it is assumed that the parser first pursues only a single rather sketchy candidate analysis on the basis of simple syntactic heuristics. The resulting syntactic skeleton is fleshed out in the following parsing stage(s), in which other types of information are furnished to yield a full semantic, contextually integrated representation. If, however, additional information conflicts with the first-stage syntactic sketch, the parser has to revise the structure in a cognitively costly second step dedicated to reanalysis. Other two-stage accounts assume that the parser already uses non-syntactic information during the first pass (Crocker, 1996; van Gompel et al., 2001; Pritchett, 1992, among others), but, crucially, parallel processing is prohibited and only a single analysis is pursued at any given parsing stage.

Since, in these models, the role of phonological information is not well accounted for, prosodic effects on syntactic structure building are not easily explained. The situation becomes worse if one considers effects of prosodic structure that is not explicitly provided in the input (as in the case of reading), especially if prosody affects the initial stages of structure building (i.e. the computation of the syntactic category of ambiguous words).

A division between first pass parsing (syntactic pre-processing) and reanalysis (integration of other information) is not assumed in the current OT model. Rather, the model adopts the notion of Stevenson and Smolensky (2006), who state that if reanalysis is understood as a change in the parser's representation of what the preferred interpretation of the input is, any addition of input necessarily constitutes some kind of reanalysis. That is, focusing on the interpretation of syntactic structure, every new piece of input requires the processing mechanism to revise the current analysis either by filling empty slots in the structure (e.g. missing heads or arguments) or by actually changing syntactic relations that had been established in previous parsing steps. The latter process is costly if the change is forced by high ranking constraints. The revision will incur only mild costs if low ranking constraints are responsible for it. That is, the degree of difficulty is determined by the change in the constraint violation pattern of the optimal candidate at word *w* relative to the previous word *w-1*. Thus, instead of two qualitatively different parsing stages as assumed in serial parsing architectures, OT envisions a single interpretation mechanism that integrates all knowledge sources for which there is relevant information in the input and, at the same time, may still account for largely differing processing costs.

4.4.2 A deterministic model for gradient data?

Even though the OT model makes clear and testable predictions of processing preferences, those predictions are rather coarse-grained. In its current format, the model is not suited for making numerical predictions but it yields an ordinal estimate of processing difficulty at most. Note also that this OT model produces deterministic parsing results. There is only one candidate structure for each parsing step that wins the competition. The parsing data are certainly gradient rather than deterministic and not all participants in the experiments behave in exactly the way the model predicts. However, the model may well reflect the overall preferences that hold for the population of participants in the experiment.

The notion of the strict ordinal constraint domination in classic OT implies a 'winner-take-all' system and thus excludes fine-grained modeling of gradient data on principled grounds (Gibson and Broihier, 1998). Some extensions of the OT framework, particularly stochastic OT (Boersma, 1998; Boersma and Hayes, 2001; Jäger, 2004), consider ranking the constraints on a continuous scale instead of the strict ordinal ranking. That way, the position of a given constraint in the hierarchy is not only defined by the relative order of constraints but also by the distance between the constraints on that scale. For the stochastic evaluation procedure, the constraint hierarchy is modified by adding normally distributed noise to each constraint position in the hierarchy such that its rank is not fixed but dwells in the area defined by the normal distribution. Depending on the proximity of the constraints on the scale, their Gaussian distributions may overlap to a certain degree. The overlap of the normal distributions then determines the probabilistic dominance relationship between the constraints and the degree to which their ranking may be reversed, yielding variable outputs.

However, applying stochastic OT to the case at hand is non-trivial. The two reading experiments that confirm the prosodic effect on syntactic ambiguity resolution have a variety of dependent variables: in the case of oral reading, the number of accentuations on the ambiguous *mehr* and the pause duration before the disambiguating phrase were chosen; for silent reading, we examined several standard reading times and fixation probabilities derived from the eye-tracking record.

Even though the results of these dependent variables are complementary, they naturally differ numerically. It is far from obvious which dependent variable should be chosen as the reference mark for the assignment of the numerical rank to the constraints and what amount of noise should be added to that value. Without a clearly defined link between a dependent variable and the higher-level linguistic processes it might reflect, the choice of the reference variable remains an arbitrary one. Also, since we seek to bring together grammar and processing as close as possible, the numerical ranking should be established independent of the processing data as well. We have good reason for the ordinal ranking, but it is unclear how to motivate the constraint hierarchy on a continuous scale without recourse to the processing data.

Unless and until a highly articulated model linking dependent variables of the experiments with higher-level linguistic processes is formulated, we abstract away from the gradience in the data and have to make do with modeling the systematic parsing preferences with a deterministic OT processor.

4.5 Conclusion

The OT account of sentence reading outlined here has several advantages over other theories: First, the OT approach makes direct use of grammatical principles to determine parsing preferences at any given parsing step. The constraints and the ranking used to model the performance data have independent support from the general competence grammar. For the kind of ambiguity studied here, the OT grammar suffices to make clear and testable processing predictions. The present model achieves this goal without making reference to extra-grammatical features of the input (such as word frequency) or working memory constraints.

Nevertheless, it might be possible to formulate extra-grammatical requirements in optimality theoretic terms that could be integrated into this framework to explain processing facts that are not reducible to core grammar. The architecture of the model is parsimonious and establishes clarity: a very small set of grammatical constraints allows the model's predictions to be derived. Still, the present OT approach allows complex, non-trivial interactions of various grammatical modules in that it integrates constraints from syntax, phonology and the respective interface. That way, it allows the modeling of prosodic effects on syntactic ambiguity resolution that cannot easily be modeled in frameworks that assume a merely unidirectional syntax-phonology interface. Moreover, the model is easily scalable; it can be flexibly adapted and enlarged to fit phenomena in which semantic or contextual constraints interact with syntactic and phonological ones. What sets it apart from other constraint-based models is its modality independence. None of the constraints used here is specifically designed for the task the model was tested on, viz. reading. The same constraint hierarchy can potentially be used to make predictions about performance in listening, speaking or writing. That is, the model answers the call for a more integrative account of language production and comprehension – a demand that has been voiced again and again in psycholinguistics (e.g. Cutler and Norris, 1999; Ferreira, 2003; Garrett, 2000; Levelt et al., 1999; Pickering and Garrod, 2007).

Certainly, fundamental issues remain unresolved: For one thing, as discussed above, it is not trivial to relate the merely ordinal expression of processing preferences derived from the OT model to numerical data from psycholinguistic experiments. In addition, it is undisputed that factors like word frequency or the limitations of working memory are important forces in sentence comprehension and production. It would be far-fetched to assume that these factors are reducible to the workings of purely grammatical constraints. At this stage, it is unclear how frequency or memory effects could be integrated in the present OT model. In spite of these desiderata, the success of this model and its predecessors suggests that the general approach of applying OT to language performance is a fruitful one. The model introduced here can be seen as an extension of existing OT parsing accounts (Fanselow et al., 1999; Hoeks and Hendriks, 2011; Stevenson and Smolensky, 2006). The model shows how cross-modular effects in sentence processing and effects of production-driven comprehension may be captured within a unified model of linguistic competence and performance.

Chapter 5

Summary / Zusammenfassung

5.1 Summary in English

The results of two reading studies presented in this work promote the idea that supralexical prosody and, particularly, linguistic rhythm are part of the mental representation constructed in the process of reading. Specifics of this representation are shown to affect sentence comprehension in various ways.

In the introduction, I argue that the preference for rhythmic alternation can be understood as a phonotactic constraint that evaluates the sequence of phonological constituents within and beyond the word level. Owing to its supralexical nature, this prosodic-phonological constraint may have a bearing on word order and, by implication, on syntactic structure building.

The first study attests that a written word sequence involving a stress clash configuration and thus forcing readers to deviate from the optimal rhythmic alternation of stressed and unstressed syllables may critically hamper the processing of complex, non-canonical sentences. At points of high memory demand, the cognitive load associated with a stress clash configuration may lead to the (temporary) overwriting of memory traces necessary for the processing of syntactic long distance dependencies. I argue that this overwriting becomes possible because the clash information implicit in the written string is processed immediately while access to stored information in working memory may be tardy.

The second study shows that the recognition of lexical-syntactic category information in ambiguous sentences may take its cue from the readers' preference for rhythmic alternation of stressed and unstressed syllables. Specifically, two reading experiments demonstrate that the resolution of a lexical-syntactic ambiguity is determined by the immediate rhythmic environment. As lexical-syntactic category information can be considered the atomic building block of syntactic structure, the experiment verifies the predictive force of linguistic rhythm and its potential guiding function for syntactic parsing in reading.

The performance data from this latter experiment are modeled as an incremental constraint satisfaction process in the framework of an optimality theoretic parsing account. Solely making use of constraints derived from competence grammar, the model is capable of capturing the data and advocates the simultaneous application of syntactic, prosodic and syntax-phonology interface constraints in incremental processing. The model predicts that, in the case of syntactic indetermination, weak prosodic constraints may decide about syntactic ambiguity resolution. The performance-compatible OT grammar integrates syntactic parsing and prosodification, hence dissolving the strict separation of language production and comprehension. At the same time the OT model endorses a bidirectional relationship between syntax and phonology in grammar.

5.2 Zusammenfassung auf Deutsch

Diese Dissertation untersucht die Rolle der Prosodie, insbesondere des linguistischen Rhythmus' für die Satzverarbeitung beim Lesen. Mein Ziel ist, die folgenden Hypothesen empirisch zu untermauern und zu verteidigen:

- Während der Verarbeitung geschriebener Sätze nutzen Leser ihr phonologisches Wissen, um eine mentale prosodisch-phonologische Repräsentation des Textes zu generieren.
- Die mentale prosodische Repräsentation wird in Übereinstimmung mit der syntaktischen Struktur des geschriebenen Textes erstellt. Beschränkungen über die Syntax–Phonologie–Schnittstelle sorgen für die Kompatibilität der syntaktischen Analyse und der (mentalen) prosodischen Wiedergabe.
- Die implizite prosodische Struktur, die Leser dem geschriebenen Text zuweisen, umfasst Parameter wie phonologische Phrasierung und Akzentuierung, aber auch untergeordnete prosodische Eigenschaften wie Wortbetonung und linguistischen Rhythmus, welcher sich aus der Abfolge von betonten und unbetonten Silben ergibt.
- Rein phonologische Wohlgeformtheitsbedingungen begleiten und beeinflussen die syntaktische Analyse von Beginn an, d.h. bereits auf der Ebene der Wortkategorieerkennung. Im Falle von syntaktischer Unterspezifizierung können Analyseentscheidungen allein auf Grundlage von phonologischen Faktoren getroffen werden.
- Beim Lesen ist die mentale, lexikalisch-prosodische Information der Wörter unter bestimmten Bedingungen früher für das Verarbeitungssystem verfügbar als höher geordnete diskursstrukturelle Information. Entsprechend können die niederschwelligen lexikalisch-phonologischen Parameter einen direkteren Einfluss auf die Satzverarbeitung haben.
- Das Satzverständnis beim Lesen ist zum Teil durch Faktoren bestimmt, die der Sprachproduktion zugeordnet werden.
- Das Zusammenspiel von syntaktischen und phonologischen Prozessen beim Lesen kann im Rahmen einer performanz-kompatiblen Kompetenzgrammatik erklärt werden.

Die Ergebnisse von zwei in diesem Werk vorgestellten Lesestudien stützen die Annahme, dass die supralexikalische Prosodie und insbesondere der linguistische Rhythmus Teil der mentalen Repäsentation sind, die beim Lesen von Sätzen generiert wird. Die Ergebnisse legen nahe, dass Details dieser Repräsentation das Satzverständnis auf unterschiedliche Weise beeinflussen.

Im Einführungskapitel schlage ich vor, dass die Präferenz für rhythmische Alternation von betonten und unbetonten Silben einer phonotaktischen Beschränkung im grammatischen Sinne entspricht, welche die phonologische Konstituentenabfolge auf Wortebene und darüber hinaus bewertet. Aufgrund ihres supralexikalischen Charakters kann diese Beschränkung Einfluss auf die Wortabfolge und folglich auf den syntaktische Strukturaufbau haben.

Die im zweiten Kapitel vorgestellte Studie zeigt, dass Wortabfolgen, die einen Gegenakzent hervorrufen und damit den Leser zwingen, von der präferierten rhythmischen Alternation abzuweichen, die Verarbeitung von syntaktisch komplexen, nicht-kanonischen Sätzen deutlich erschweren. Die durch Gegenakzent hervorgerufene kognitive Beanspruchung kann an Stellen hoher syntaktischer Komplexität zum Überschreiben von Arbeitsgedächtnisinhalten führen, welche für die Verarbeitung von syntaktischen Abhängigkeiten unabdingbar sind. Das mögliche Ergebnis dieses Prozesses ist eine Fehlinterpretation des Satzes. Der Vorgang wird dadurch erklärbar, dass die für das Überschreiben verantwortliche Information über den Gegenakzent unmittelbar verarbeitet wird, während der Zugriff auf die im Abeitsgedächtnis gespeichrte Information relativ langsam ist.

Eine zweite Lesestudie zeigt, dass die Identifikation der lexikalisch-syntaktischen Kategorie ambiger Wörter im Satzkontext durch die Erwartung rhythmischer Alternanz von betonten und unbetonten Silben beeinflusst ist. Insbesondere belegen die zwei Experimente, dass die Ambiguitätsauflösung durch das unmittelbare rhythmische Umfeld bestimmt ist. Da die lexikalisch-syntaktische Kategorie als kleinstes Element des syntaktischen Strukturaufbaus gilt, sind diese Ergebnisse als Beleg für die prädiktive Rolle der Prosodie und ihre leitende Funktion bei der syntaktischen Analyse schriftlicher Sätze zu bewerten.

Die Daten der letztgenannten Studie werden im Rahmen eines inkrementell organisierten optimalitätstheoretischen Verarbeitungsmodells abgebildet. Das Modell nutzt dafür ausschliesslich Beschränkungen, die der Kompetenzgrammatik entlehnt sind. Die Datenmuster der Leseexperimente werden durch das simultane Zusammenspiel von syntaktischen und prosodischen Beschränkungen sowie Regularitäten der Syntax–Phonologie Schnittstelle modelliert. Die Ergebnisse dieser Interaktion erlauben die Vorhersage, dass im Falle syntaktischer Unterdeterminiertheit rein prosodische Beschränkungen über die Auflösung syntaktischer Strukturambiguitäten entscheiden. Das Modell integriert syntaktische Strukturanalyse und Prosodiegenerierung in einem einzigen Prozess und löst damit die strikte Trennung von Sprachperzeption und -produktion auf. Darüber hinaus unterstreicht das Modell die Vorstellung einer bidirektionalen Beziehung zwischen Syntax und Phonologie in Grammatik und Sprachverarbeitung.

Bibliography

Abney, S., 1987. The English NP in its sentential aspect. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.

Alber, B., 2005. Clash, lapse and directionality. Natural Language and Linguistic Theory 23 (3), 485–542.

Anttila, A., Adams, M., Speriosu, M., 2010. The role of prosody in the English dative alternation. Language and Cognitive Processes 25 (7-9), 946–981.

Archangeli, D., 1997. Optimality theory: An introduction to linguistics in the 1990s. In: Archangeli, D., Langendoen, D. (Eds.), Optimality theory: An overview. Wiley-Blackwell, pp. 1–33.

Ashby, J., Clifton, C., 2005. The prosodic property of lexical stress affects eye movements during silent reading. Cognition 96 (3), B89–B100.

Ashby, J., Martin, A., 2008. Prosodic phonological representations early in visual word recognition. Journal of Experimental Psychology: Human Perception and Performance 34 (1), 224–236.

Ashby, J., Rayner, K., 2004. Representing syllable information during silent reading: Evidence from eye movements. Language and Cognitive Processes 19 (3), 391–426.

Augurzky, P., 2006. Attaching relative clauses in German: The role of implicit and explicit prosody in sentence processing. Ph.D. thesis, MPI for Human Cognitive and Brain Sciences, Leipzig.

Bader, M., 1996a. Prosodic effects and the distinction between primary and non-primary phrases. In: Proceedings of the Conference on Architectures and Mechanisms of Language Processing (AMLaP). University of Torino.

Bader, M., 1996b. Sprachverstehen: Syntax und Prosodie beim Lesen. Westdeutscher Verlag.

Bader, M., 1998. Prosodic influences on reading syntactically ambiguous sentences. In: Fodor, J., Ferreira, F. (Eds.), Reanalysis in sentence processing. Dordrecht: Kluwer, pp. 1–46.

Bates, D., Sarkar, D., 2007. lme4: Linear mixed-effects models using S4 classes. R package version 0.9975-11.

Baumann, S., Grice, M., Steindamm, S., 2006. Prosodic marking of focus domains – categorical or gradient. In: Proceedings of Speech Prosody. Dresden, pp. 301–304.

Beauvois, M., Dérouesné, J., 1979. Phonological alexia: three dissociations. Journal of Neurology, Neurosurgery and Psychiatry 42 (12), 1115–1124.

Beaver, D., Lee, H., 2004. Input-output mismatches in OT. In: Blutner, R., Zeevat, H. (Eds.), Optimality theory and pragmatics. Palgrave McMillan, pp. 112–153.

Beckman, M., Pierrehumbert, J., 1986a. Intonational structure in Japanese and English. Phonology Yearbook 3 (1), 255–309.

Beckman, M., Pierrehumbert, J., 1986b. Intonational structure in Japanese and English. Phonology Yearbook 3, 255–309.

Berg, T., 2008. Emphatic stress shift in German. Zeitschrift für Sprachwissenschaft 27 (2), 165–187.

Bergmann, A., Armstrong, M., Maday, K., 2008. Relative clause attachment in English and Spanish: A production study. In: Proceedings of the 4th conference on Speech Prosody. Campinas, Brazil, pp. 505–508.

Bierwisch, M., 1966. Regeln für die Intonation deutscher Sätze. Studia Grammatica VII, 99–201.

Bock, K., Griffin, Z., 2000. The persistence of structural priming: Transient activation or implicit learning? Journal of Experimental Psychology: General 129 (2), 177–192.

Boersma, P., 1998. Functional phonology: Formalizing the interactions between articulatory and perceptual drives. Ph.D. thesis, University of Amsterdam, The Hague.

Boersma, P., Hayes, B., 2001. Empirical tests of the gradual learning algorithm. Linguistic Inquiry 32 (1), 45–86.

Bohn, K., Wiese, R., Domahs, U., 2011. The status of the Rhythm Rule within and across word boundaries in German. In: Proceedings of ICPhS XVII.

Breen, M., Clifton, C., 2011. Stress matters: Effects of anticipated lexical stress on silent reading. Journal of Memory and Language 64 (2), 153–170.

Carlson, K., 2001. The effects of parallelism and prosody in the processing of gapping structures. Language and Speech 44 (1), 1–26.

Chafe, W., 1988. Punctuation and the prosody of written language. Written Communication 5 (4), 395–426.

Chomsky, N., 1965. Aspects of the Theory of Syntax. MIT Press, Cambridge, MA.

Chomsky, N., 1981. Lectures on Government and Binding. Foris, Dordrecht.

Chomsky, N., 1995. The Minimalist Program. MIT Press, Cambridge, MA.

Christiansen, M., Chater, N., 2001. Connectionist psycholinguistics: Capturing the empirical data. Trends in Cognitive Sciences 5 (2), 82–88.

Clifton, C., Staub, A., Rayner, K., 2007. Eye movements in reading words and sentences. In: van Gompel, R., Fischer, M., Murray, W., Hill, R. L. (Eds.), Eye movements: A window on mind and brain. Elsevier, Amsterdam, Ch. 15, pp. 341–372.

Crocker, M., 1996. Computational psycholinguistics: An interdisciplinary approach to the study of language. Kluwer, Dordrecht.

Cutler, A., Dahan, D., Van Donselaar, W., 1997. Prosody in the comprehension of spoken language: a literature review. Language and Speech 40, 141–201.

Cutler, A., Norris, D., 1999. Sharpening Ockham's razor. Behavioral and Brain Sciences 22 (01), 40–41.

Dell, G., Chang, F., Griffin, Z., 1999. Connectionist models of language production: Lexical access and grammatical encoding. Cognitive Science 23 (4), 517–542.

Dilley, L., McAuley, J., 2008. Distal prosodic context affects word segmentation and lexical processing. Journal of Memory and Language 59 (3), 294–311.

Eady, S., Cooper, W., Klouda, G., Mueller, P., Lotts, D., 1986. Acoustical characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments. Language and Speech 29 (3), 233.

Fanselow, G., Schlesewsky, M., Cavar, D., Kliegl, R., 1999. Optimal parsing, syntactic parsing preferences, and optimality theory, Rutgers Optimality Archive, ROA-367.

Ferreira, F., 2003. The misinterpretation of noncanonical sentences. Cognitive Psychology 47 (2), 164–203.

Féry, C., Hartmann, K., 2005. Focus and prosodic structures of German Gapping and Right Node Raising. The Linguistic Review 22, 67–114.

Féry, C., Herbst, L., 2004. German sentence accent revisited. ISIS Working Papers of the SFB 632 1, 43–75.

Féry, C., Kügler, F., 2008. Pitch accent scaling on given, new and focused constituents in German. Journal of Phonetics 36 (4), 680–703.

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

Fodor, J., 1998. Learning to parse? Journal of Psycholinguistic Research 27 (2), 285–319.

Fodor, J. D., 2002. Psycholinguistics cannot escape prosody. In: Proceedings of the 1st International Conference on Speech Prosody. Aix-en-Provence, pp. 83–88.

Forster, K., Forster, J., 2003. DMDX: A Windows display program with millisecond accuracy. Behavior Research Methods, Instruments and Computers 35 (1), 116–124.

Frazier, L., 1987. Sentence processing: A tutorial review. In: Coltheart, M. (Ed.), Attention and performance XII. Lawrence Erlbaum Associates, pp. 559–586.

Frazier, L., Carlson, K., Clifton, C., 2006. Prosodic phrasing is central to language comprehension. Trends in Cognitive Science 10 (6), 244–249.

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.

Frost, R., 1998. Toward a strong phonological theory of visual word recognition: True issues and false trails. Psychological Bulletin 123, 71–99.

Fuchs, L., Fuchs, D., Hosp, M., Jenkins, J., 2001. Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific Studies of Reading 5 (3), 239–256.

Garnsey, S., Pearlmutter, N., Myers, E., Lotocky, M., 1997. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. Journal of Memory and Language 37 (1), 58–93.

Garrett, M., 1975. The Analysis of Sentence Production1. In: Bower, G. (Ed.), The psychology of learning and motivation. Academic Press, pp. 133–177.

Garrett, M., 2000. Remarks on the architecture of language processing systems. In: Grodzinsky, Y., Shapiro, L., Swinney, D. (Eds.), Language and the brain: Representation and processing. Academic Press, pp. 31–69.

Gelman, A., Hill, J., 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge.

Ghini, M., 1993. Φ-formation in Italian: A new proposal. Toronto Working Papers in Linguistics, 41–78.

Gibson, E., 2000. Dependency locality theory: A distance-based theory of linguistic complexity. In: Marantz, A., Miyashita, Y., O'Neil, W. (Eds.), Image, Language, Brain: Papers from the 1st Mind Articulation Project Symposium. MIT Press, Cambridge, MA, pp. 95–126.

Gibson, E., Broihier, K., 1998. Optimality theory and human sentence processing. In: Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M., Pesetsky, D. (Eds.), Is the best good enough? MIT Press, pp. 157–191.

Gibson, E., Thomas, J., 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. Language and Cognitive Processes 14(3), 225–248.

Giegerich, H., 1985. Metrical phonology and phonological structure. German and English. Cambridge University Press, Cambridge.

Golston, C., Riad, T., 2000. The phonology of Classical Greek meter. Linguistics 38, 99–168.

Grabe, E., Warren, P., 1995. Stress Shift: Do speakers do it or listeners hear it. In: Connell, B., Arvaniti, A. (Eds.), Phonology and phonetic evidence. Cambridge University Press, pp. 95–109.

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., 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. Focus, mode and the nucleus. Journal of Linguistics 19 (2), 377–417.

Haider, H., 2011. Grammatische Illusionen – lokal wohlgeformt – global deviant. Zeitschrift für Sprachwissenschaft 30 (2), 223–257.

Hartmann, K., 2000. Right node raising and gapping: Interface conditions on prosodic deletion. John Benjamins Publishing Company.

Hawkins, J. A., 1994. A Performance Theory of Order and Constituency. Cambridge University Press, New York.

Hayes, B., 1989. The prosodic hierarchy in meter. In: Kiparsky, P., Youmans, G. (Eds.), Rhythm and Meter. Academic Press, Orlando, pp. 201–260.

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

Hendriks, P., De Hoop, H., 2001. Optimality theoretic semantics. Linguistics and Philosophy 24 (1), 1–32.

Hirose, Y., 2003. Recycling prosodic boundaries. Journal of Psycholinguistic Research 32 (2), 167–195.

Hoeks, J., Hendriks, P., 2011. Optimality theory and human sentence processing: Towards a cross-modular analysis of coordination. Journal of Cognitive Science 12, 83–128.

Hwang, H., Schafer, A., 2009. Constituent length affects prosody and processing for a dative NP ambiguity in Korean. Journal of Psycholinguistic Research 38 (2), 151–175.

Hwang, H., Steinhauer, K., 2011. Phrase length matters: The interplay between implicit prosody and syntax in Korean garden path sentences. Journal of Cognitive Neuroscience 23 (11), 3555–75.

Inkelas, S., Zec, D., 1995. The phonology-syntax interface. In: Goldsmith, J. (Ed.), The handbook of phonological theory. Blackwell, pp. 535–549.

Ito, K., Speer, S., 2008. Anticipatory effects of intonation: Eye movements during instructed visual search. Journal of Memory and Language 58 (2), 541–573.

Jackendoff, R., 2003. Foundations of language: Brain, meaning, grammar, evolution. Oxford University Press.

Jäger, G., 2004. Learning constraint sub-hierarchies. The bidirectional gradual learning algorithm. In: Blutner, R., Zeevat, H. (Eds.), Optimality Theory and pragmatics. Palgrave McMillan, pp. 251–287.

Joshi, A. K., 1990. Processing crossed and nested dependencies: An automaton perspective on the psycholinguistic results. Language and Cognitive Processes 5(1), 1–27.

Jun, S., 2003. Prosodic phrasing and attachment preferences. Journal of Psycholinguistic Research 32 (2), 219–249.

Jun, S., 2010. The implicit prosody hypothesis and overt prosody in English. Language and Cognitive Processes 25 (7), 1201–1233.

Kager, R., 1989. A metrical theory of stress and destressing in English and Dutch. Walter de Gruyter.

Kager, R., 1999. Optimality Theory. Cambridge University Press.

Kelly, M., Bock, J., 1988. Stress in time. Journal of Experimental Psychology: Human Perception and Performance 14 (3), 389–403.

Kennedy, A., Pynte, J., 2005. Parafoveal-on-foveal effects in normal reading. Vision Research 45 (2), 153–168.

Kentner, G., 2007. Prosodic Phrasing in Elliptic and Non-elliptic Coordinations. In: Proceedings of the XVI. International Congress of Phonetic Sciences. Saarbrücken, pp. 1125–1128.

Kentner, G., 2012. Linguistic rhythm guides parsing decisions in written sentence comprehension. Cognition 123 (1), 1–20.

Kentner, G., Féry, C., Alter, K., 2008. Prosody in Speech Production and Perception: The Case of Right Node Raising in English. In: Steube, A. (Ed.), The Discourse Potential of Underspecified Structure. Mouton de Gruyter, Berlin, pp. 207–223.

Kiparsky, P., 1966. Über den deutschen Akzent. Studia Grammatica VII, 69–98.

Kjelgaard, M., Speer, S., 1999. Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. Journal of Memory and Language 40 (2), 153–194.

Kleiman, G., 1975. Speech recoding in reading. Journal of Verbal Learning and Verbal Behavior 14 (4), 323–339.

Kliegl, R., Engbert, R., 2005. Fixation durations before word skipping in reading. Psychonomic Bulletin and Review 12 (1), 132–138.

Kliegl, R., Nuthmann, A., Engbert, R., 2006. Tracking the mind during reading: The influence of past, present, and future words on fixation durations. Journal of Experimental Psychology: General 135, 12–35.

Kondo, T., Mazuka, R., 1996. Prosodic planning while reading aloud: On-line examination of Japanese sentences. Journal of Psycholinguistic Research 25 (2), 357–381.

Koriat, A., Greenberg, S., Kreiner, H., 2002. The extraction of structure during reading: Evidence from reading prosody. Memory and Cognition 30 (2), 270–280.

Kreiner, H., Koriat, A., 2005. The Role of Reading Prosody in Syntactic and Semantic Integration: Evidence from Eye Movements. In: Discourse and Prosody as a complex interface. Aix-en-Provence.

Lamb, S., 1998. Pathways of the brain: The neurocognitive basis of language. John Benjamins, Amsterdam.

Lechner, W., 2003. Comparatives and DP-structure. Glot International 7 (3), 1–9.

Levelt, W., Roelofs, A., Meyer, A., 1999. A theory of lexical access in speech production. Behavioral and Brain Sciences 22 (01), 1–38.

Levin, H., Addis, A., 1979. The eye-voice span. MIT Press, Cambridge.

Liberman, M., Prince, A., 1977. On stress and linguistic rhythm. Linguistic Inquiry 8 (2), 249–336.

Logačev, P., Vasishth, S., 2006. The em package for computing eyetracking measures. University of Potsdam, Potsdam, Germany.

MacDonald, M. C., Perlmutter, N. J., Seidenberg, M. S., 1994. The lexical nature of syntactic ambiguity resolution. Psychological Review 101(4), 676–703.

Marslen-Wilson, W., Tyler, L., Warren, P., Grenier, P., Lee, C., 1992. Prosodic effects in minimal attachment. The Quarterly Journal of Experimental Psychology Section A 45 (1), 73–87.

McCarthy, J., 2002. A thematic guide to Optimality Theory. Cambridge University Press.

McClelland, J., Elman, J., 1986. The TRACE model of speech perception. Cognitive Psychology 18 (1), 1–86.

McCutchen, D., Perfetti, C., 1982. The visual tongue-twister effect: Phonological activation in silent reading. Journal of Verbal Learning and Verbal Behavior 21 (6), 672–687.

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

Nespor, M., Vogel, I., 2007. Prosodic phonology. Mouton de Gruyter, Berlin.

Niebuhr, O., 2009. F0-based rhythm effects on the perception of local syllable prominence. Phonetica 66 (1–2), 95–112.

Phillips, C., 1996. Order and structure. Ph.D. thesis, MIT.

Pickering, M., Garrod, S., 2007. Do people use language production to make predictions during comprehension? Trends in Cognitive Sciences 11 (3), 105–110.

Pickering, M., van Gompel, R., 2006. Syntactic parsing. In: Traxler, M. (Ed.), Handbook of psycholinguistics. Vol. 2. Academic Press New York, pp. 455–503.

Prince, A., 1983. Relating to the grid. Linguistic Inquiry 14 (1), 19–100.

Prince, A., Smolensky, P., 1993. Optimality theory: Constraint interaction in generative grammar. University Center for Cognitive Science, Piscataway, Rutgers: NJ.

Prince, A., Smolensky, P., 1997. Optimality: From neural networks to universal grammar. Science 275 (5306), 1604–1610.

Prince, A., Smolensky, P., 2004. Optimality Theory: Constraint interaction in generative grammar. In: McCarthy, J. (Ed.), Optimality theory in phonology: A reader. Wiley-Blackwell, pp. 3–71.

Pritchett, B. L., 1992. Grammatical competence and parsing performance. University of Chicago Press, Chicago.

Pullum, G., Zwicky, A., 1988. The syntax-phonology interface. In: Newmeyer, F. J. (Ed.), Linguistics: The Cambridge Survey. Vol. 1. Cambridge University Press, pp. 255–280.

Pynte, J., Prieur, B., 1996. Prosodic breaks and attachment decisions in sentence processing. Language and Cognitive Processes 11, 165–192.

Quené, H., van den Bergh, H., 2004. On multi-level modeling of data from repeated measures designs: A tutorial. Speech Communication 43 (1), 103–122.

Rastle, K., Coltheart, M., 1998. Whammies and double whammies: The effect of length on nonword reading. Psychonomic Bulletin and Review 5 (2), 277–282.

Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin 124 (3), 372–422.

Rayner, K., Kambe, G., Duffy, S., 2000. The effect of clause wrap-up on eye movements during reading. The Quarterly Journal of Experimental Psychology Section A 53 (4), 1061–1080.

Reinhart, T., 2006. Interface strategies – optimal and costly computations. MIT Press.

Rice, C., Svenonius, P., 1998. Prosodic V2 in Northern Norwegian. Manuscript, University of Tromsø.

Saffran, E., Marin, O., 1977. Reading without phonology: Evidence from aphasia. The Quarterly Journal of Experimental Psychology 29 (3), 515–525.

Sag, I., Wasow, T., 2011. Performance-compatible competence grammar. In: Borsley, R. D., Börjars, K. (Eds.), Nontransformational Syntax: Formal and explicit models of grammar. Wiley-Blackwell, Oxford, pp. 359–377.

Samek-Lodovici, V., 2005. Prosody–syntax interaction in the expression of focus. Natural Language and Linguistic Theory 23 (3), 687–755.

Sanford, A., Sturt, P., 2002. Depth of processing in language comprehension: Not noticing the evidence. Trends in Cognitive Sciences 6 (9), 382–386.

Schafer, A., Carlson, K., Clifton jr, C., Frazier, L., 2000. Focus and the interpretation of pitch accent: Disambiguating embedded questions. Language and Speech 43 (1), 75–105.

Schlüter, J., 2003. Phonological determinants of grammatical variation in English: Chomsky's worst possible case. In: Rohdenburg, G., Mondorf, B. (Eds.), Determinants of grammatical variation in English. Mouton De Gruyter, Berlin, pp. 69–118.

Schlüter, J., 2005. Rhythmic grammar: the influence of rhythm on grammatical variation and change in English. Mouton de Gruyter, Berlin.

Schmidt-Kassow, M., Kotz, S., 2009. Attention and perceptual regularity in speech. NeuroReport 20 (18), 1643–1647.

Schwanenflugel, P., Hamilton, A., Kuhn, M., Wisenbaker, J., Stahl, S., 2004. Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. Journal of Educational Psychology 96 (1), 119–129.

Seidenberg, M., Plaut, D., 1998. Evaluating word-reading models at the item level: Matching the grain of theory and data. Psychological Science 9 (3), 234–237.

Selkirk, E., 1984. Phonology and syntax. MIT Press Cambridge, Mass.

Selkirk, E., 1995. Sentence prosody: Intonation, stress, and phrasing. In: Goldsmith, J. (Ed.), The handbook of phonological theory. Blackwell, Oxford, pp. 550–569.

Selkirk, E., 2002. Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in English. In: Proceedings of Speech Prosody 2002. Aix-en-Provence, pp. 643–646.

Shallice, T., Warrington, E., 1975. Word recognition in a phonemic dyslexic patient. The Quarterly Journal of Experimental Psychology 27 (2), 187–199.

Shih, S., Grafmiller, J., Futrell, R., Bresnan, J., 2009. Rhythm's role in genitive and dative construction choice in spoken English. In: 31st annual meeting of the Linguistics Association of Germany (DGfS), University of Osnabrück, Germany.

Slowiaczek, M., Clifton, C., et al., 1980. Subvocalization and reading for meaning. Journal of Verbal Learning and Verbal Behavior 19 (5), 573–582.

Smolensky, P., 1996. On the comprehension/production dilemma in child language. Linguistic Inquiry 27 (4), 720–731.

Smolensky, P., Legendre, G., 2006. The harmonic mind. MIT Press.

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.

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.

Speyer, A., 2008. Topicalization and clash avoidance. On the interaction of prosody and syntax in the history of English with a few glimpses at German. Ph.D. thesis, University of Pennsylvania.

Speyer, A., 2010. Topicalization and stress clash avoidance in the history of English. Mouton De Gruyter, Berlin.

Steinhauer, K., 2003. Electrophysiological correlates of prosody and punctuation. Brain and Language 86 (1), 142–164.

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

Steinhauer, K., Friederici, A., 2001. Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. Journal of Psycholinguistic Research 30 (3), 267–295.

Stevenson, S., Smolensky, P., 2006. Optimality in sentence processing. In: Smolensky, P., Legendre, G. (Eds.), The harmonic mind. Vol. 2. MIT Press, pp. 307–338.

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

Swets, B., Desmet, T., Hambrick, D., Ferreira, F., 2007. The role of working memory in syntactic ambiguity resolution: A psychometric approach. Journal of Experimental Psychology: General 136 (1), 64–81.

Tabor, W., Juliano, C., Tanenhaus, M., 1997. Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. Language and Cognitive Processes 12 (2), 211–271.

Tabor, W., Tanenhaus, M., 1999. Dynamical models of sentence processing. Cognitive Science 23 (4), 491–515.

Truckenbrodt, H., 1995. Phonological phrases: Their relation to syntax, focus, and prominence. Ph.D. thesis, Massachusetts Institute of Technology.

Truckenbrodt, H., 2006. Phrasal stress. In: Brown, K. (Ed.), The encyclopedia of languages and linguistics, 2nd Edition. Vol. 9. Elsevier, Amsterdam, pp. 572–579.

Truckenbrodt, H., 2007. The syntax-phonology interface. In: de Lacy, P. (Ed.), The Cambridge handbook of phonology. Cambridge University Press, Cambridge, pp. 435–456.

Truckenbrodt, H., Darcy, I., 2010. Object clauses, movement, and phrasal stress. In: Erteshik-Shir, N., Rochman, L. (Eds.), The Sound Patterns of Syntax. Oxford University Press, Oxford, pp. 189–216.

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.

van Casteren, M., Davis, M., 2006. Mix, a program for pseudorandomization. Behavior Research Methods 38 (4), 584–589.

van Gompel, R., Pickering, M., Traxler, M., 2001. Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. Journal of Memory and Language 45 (2), 225–258.

Visch, E., 1999. The rhythmic organization of compounds and phrases. In: van der Hulst, H. (Ed.), Word prosodic systems in the languages of Europe. Walter de Gruyter, pp. 161–231.

Wagner, M., Watson, D., 2010. Experimental and theoretical advances in prosody: A review. Language and Cognitive Processes 25, 1–41.

Warren, P., Grabe, E., Nolan, F., 1995. Prosody, phonology and parsing in closure ambiguities. Language and Cognitive Processes 10 (5), 457–486.

Watt, S., Murray, W., 1996. Prosodic form and parsing commitments. Journal of Psycholinguistic Research 25 (2), 291–318.

Wheeldon, L., 2000. Generating prosodic structure. In: Wheeldon, L. (Ed.), Aspects of language production. Psychology Press, pp. 249–274.

Wiese, R., 2000. The phonology of German. Oxford University Press.

Wilder, C., 1997. Some properties of ellipsis in coordination. In: Alexiadou, A., Hall, T. (Eds.), Studies on universal grammar and typological variation. John Benjamins, Amsterdam, pp. 59–107.

Wilshire, C., 1999. The tongue twister paradigm as a technique for studying phonological encoding. Language and Speech 42 (1), 57.

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

Zubizarreta, M., 1998. Prosody, focus, and word order. MIT Press.

Appendix

Sentences used as items in the experiment of chapter 2

1. Der Maurer {holt, lächelt} und {der Admiral, die Lehrerin} isst {Kuchen, Gebäck} mit Marzipan.
2. Die Soldatin {holt, lächelt} und {der General, der Hausmeister} testet {Fahrräder, Gewehre} aus Holland.
3. Der Reiseveranstalter {mietet, schwitzt} und {die Agentur, der Arbeiter} putzt {Ferienhäuser, Versammlungsräume} auf Usedom.
4. Die Studentin {mietet, schwitzt} und {der Assistent, der Busfahrer} steuert {Reisebusse, Geländewagen} von BMW.
5. Der Angestellte {meldet, redet} und {der Fabrikant, der Bauleiter} prüft {Ausfälle, Verluste} durch Diebstähle.
6. Der Zeuge {meldet, redet} und {der Detektiv, der Aufseher} sucht {Tatverdächtige, Verdächtige} im Keller.
7. Der Dozent {kauft, lacht} und {der Gitarrist, der Handwerker} raubt {Ölgemälde, Gemälde} von Picasso.
8. Der Kunde {kauft, lacht} und {der Amateur, der Hersteller} bringt {Stahlträger, Gerüstteile} aus Japan.
9. Der Hotelgast {ordnet, wandert} und {der Kapitän, der Optiker} stempelt {Unterlagen, Behördenbriefe} für Bedürftige.
10. Der Rentner {ordnet, wandert} und {der Musikant, der Musiker} sammelt {Schallplatten, Gerümpel} der Beatles.
11. Der Mönch {sammelt, schweigt} und {der Kardinal, der Prediger} kauft {Goldmünzen, Gebrauchtwagen} aus Irland.
12. Der Dorfpolizist {sammelt, schweigt} und {der Journalist, der Buchhalter} ordnet {Anzeigen, Verlustanzeigen} wegen Diebstahls.
13. Der Lehrling {stapelt, streikt} und {der Philosoph, der Ausbilder} streicht {Rahmen, Verschalungen} aus Holz.
14. Der Hersteller {stapelt, streikt} und {der Germanist, der Pfadfinder} lackiert {Türschilder, Gewinde} für Stammkunden.

15. Der Gewerkschafter {fordert, nickt} und {die Agentur, der Botschafter} erhält {Sonderzahlungen, Verzehrgutscheine} auf Vertrauensbasis.

16. Der Minister {fordert, nickt} und {das Tribunal, der Machthaber} beschlie§t {Lohnerhöhungen, Gesetzesänderungen} für Werftarbeiter.

17. Der Fabrikarbeiter {säubert, jammert} und {der Archivar, die Künstlerin} schmuggelt {Porzellanvasen, Gefässe} mit Henkel.

18. Der Hilfsarbeiter {säubert, jammert} und {der Diplomat, der Botschafter} liefert {Sanduhren, Geräte} aus England.

19. Die Sängerin {lobt, weint} und {der Pianist, die Königin} ehrt {Musiker, Gewinner} des Wettbewerbs.

20. Die Vorsitzende {lobt, weint} und {das Dekanat, der Manager} fördert {Bürgerinitiativen, Verantwortliche} aus Norwegen.

21. Der Zirkusartist {liebt, kämpft} und {der Germanist, der Zuschauer} schreibt {Briefe, Gedichte} in Schönschrift.

22. Die Boxerin {liebt, kämpft} und {der Fotograf, der Kritiker} liest {Bücher, Erzählungen} von Kleist.

23. Die Ingenieurin {plant, schläft} und {die Gärtnerei, der Holzhändler} bepflanzt {Parkanlagen, Gewerbegebiete} in Bremen.

24. Der Bürgermeister {plant, schläft} und {der Kommandant, die Künstlerin} filmt {Probebohrungen, Beschlagnahmungen} im Niemandsland.

25. Der Küchenchef {schneidet, flüstert} und {der Astronaut, der Kundschafter} probiert {Truthahn, Gemüse} mit Füllung.

26. Der Kochlehrling {schneidet, flüstert} und {der Absolvent, die Kellnerin} serviert {Lendenbraten, Geschnetzeltes} vom Schwein.

27. Der Händler {druckt, grinst} und {der Demokrat, der Komiker} verteilt {Flugblätter, Beschwerdebriefe} gegen Terroristen.

28. Der Direktor {druckt, grinst} und {der Monarchist, der Inhaber} sortiert {Geldscheine, Verträge} für Bankkunden.

Sentences used as items in the experiments of chapter 3

1ab Der Physiker glaubt, dass man rechnerisch nicht mehr {nachweisen, ermitteln} kann, ob es einen zehnten Planeten gibt.

1cd Der Physiker glaubt, dass man rechnerisch nicht mehr {nachweisen, ermitteln} kann, als das Gewicht des Körpers.

2ab Karlo sagte, dass er am Tatort nicht mehr {nachweisen, ermitteln} konnte, ob die DNA übereinstimmt.

2cd Karlo sagte, dass er am Tatort nicht mehr {nachweisen, ermitteln} konnte, als die ungefähre Tatzeit.

3ab Anton Müller denkt, dass der Direktor nicht mehr {hinnehmen, gestatten} sollte, dass die Zufahrt ständig zugeparkt ist.

3cd Anton Müller denkt, dass der Direktor nicht mehr {hinnehmen, gestatten} sollte, als einen Nachtdienst pro Woche.

4ab Rita denkt, dass man als Chefin nicht mehr {hinnehmen, gestatten} sollte, dass dauernd Überstunden anfallen.

4cd Rita denkt, dass man als Chefin nicht mehr {hinnehmen, gestatten} sollte, als einen Tag Sonderurlaub im Monat.

5ab Tim meint, dass man den Lehrern nicht mehr {anbieten, versprechen} sollte, auf Schokolade ganz zu verzichten.

5cd Tim meint, dass man den Lehrern nicht mehr {anbieten, versprechen} sollte, als das Erledigen der Hausaufgaben.

6ab Uta Wendt meint, dass Mediziner prinzipiell nicht mehr {anbieten, versprechen} sollten, jeden Patienten zu behandeln.

6cd Uta Wendt meint, dass Mediziner prinzipiell nicht mehr {anbieten, versprechen} sollten, als sie selbst garantieren können.

7ab Andreas erzählte, dass Paul letzten Mittwoch nicht mehr {darlegen, bezeugen} wollte, was er am Montag gesehen hatte.

7cd Andreas erzählte, dass Paul letzten Mittwoch nicht mehr {darlegen, bezeugen} wollte, als sowieso allseits bekannt war.

8ab Wolfgang weiß, dass Ulf vor Gericht nicht mehr {darlegen, bezeugen} wollte, wie gefährlich die Arbeit ist.

8cd Wolfgang weiß, dass Ulf vor Gericht nicht mehr {darlegen, bezeugen} wollte, als die Polizei bereits wusste.

9ab Joachim beklagt, dass Karola am Donnerstag nicht mehr {zugeben, gestehen} wollte, dass sie Raucherin ist.

9cd Joachim beklagt, dass Karola am Donnerstag nicht mehr {zugeben, gestehen} wollte, als ihre Abhängigkeit von Nikotin.

10ab Rufus empfiehlt, dass man vor Gericht nicht mehr {zugeben, gestehen} sollte, dass man schuldig ist.

10cd Rufus empfiehlt, dass man vor Gericht nicht mehr {zugeben, gestehen} sollte, als bereits bewiesen ist.

11ab Jeder wusste, dass Martin der Lehrerin nicht mehr {mitteilen, beschreiben} wollte, was am Hafen passiert ist.

11cd Jeder wusste, dass Martin der Lehrerin nicht mehr {mitteilen, beschreiben} wollte, als die Planung der Abschlussfeier.

12ab Nina befürchtet, dass Johannes dem Professor nicht mehr {mitteilen, beschreiben} will, wie es in der Mensa zugeht.

12cd Nina befürchtet, dass Johannes dem Professor nicht mehr {mitteilen, beschreiben} will, als auf der Sitzung der Verwaltung.

13ab Jan fragte, warum die Minister gestern nicht mehr {vorschlagen, besprechen} wollten, die Ortsumgehung zu bauen.

13cd Jan fragte, warum die Minister gestern nicht mehr {vorschlagen, besprechen} wollten, als die Preisverleihung an Biermann.

14ab Wiebke überlegte, warum die Sänger gestern nicht mehr {vorschlagen, besprechen} wollten, einen neuen Dirigenten anzuwerben.

14cd Wiebke überlegte, warum die Sänger gestern nicht mehr {vorschlagen, besprechen} wollten, als die Feier für den Dirigenten.

15ab Franziska bedauert, dass Rolf den Schülern nicht mehr {antworten, erläutern} konnte, wie die Mikrowelle funktioniert.

15cd Franziska bedauert, dass Rolf den Schülern nicht mehr {antworten, erläutern} konnte, als einen kurzen Satz.

16ab Ralf bedauert, dass Matthias den Journalisten nicht mehr {antworten, erläutern} wollte, wie er zu dem Urteil kam.

16cd Ralf bedauert, dass Matthias den Journalisten nicht mehr {antworten, erläutern} wollte, als schon in dem Urteil stand.

17ab Hans Riemers findet, dass der Boss nicht mehr {zulassen, erlauben} sollte, dass die Arbeiter dauernd Pause machen.

17cd Hans Riemers findet, dass der Boss nicht mehr {zulassen, erlauben} sollte, als die Firma sich leisten kann.

18ab Marco Schmidt findet, dass der Pfarrer nicht mehr {zulassen, erlauben} sollte, dass Touristen in der Kirche fotografieren.

18cd Marco Schmidt findet, dass der Pfarrer nicht mehr {zulassen, erlauben} sollte, als das Fotografieren ohne Blitz.

19ab Maria denkt, dass Jochen den Soldaten nicht mehr {anordnen, befehlen} sollte, die Sandsäcke hinterm Deich abzuladen.

19cd Maria denkt, dass Jochen den Soldaten nicht mehr {anordnen, befehlen} sollte, als der Truppe zuzumuten ist.

20ab Sonja denkt, dass Stefan seinen Mitarbeitern nicht mehr {anordnen, befehlen} sollte, wie die Arbeit zu erledigen ist.

20cd Sonja denkt, dass Stefan seinen Mitarbeitern nicht mehr {anordnen, befehlen} sollte, als ihre Kräfte zulassen.

21ab Der Polizist sagte, dass die Ermittler nicht mehr {feststellen, erfahren} konnten, wieviel Geld gestohlen wurde.

21cd Der Polizist sagte, dass die Ermittler nicht mehr {feststellen, erfahren} konnten, als die Haarfarbe des Täters.

22ab Rainer sagte, dass man mit Filmaufnahmen nicht mehr {feststellen, erfahren} kann, wer den Unfall verursacht hat.

22cd Rainer sagte, dass man mit Filmaufnahmen nicht mehr {feststellen, erfahren} kann, als das Geschlecht des Opfers.

23ab Der Chemiker ist sauer, weil Paula nicht mehr {ausrechnen, berechnen} wollte, wieviel Wasserstoff im Reagenzglas ist.

23cd Der Chemiker ist sauer, weil Paula nicht mehr {ausrechnen, berechnen} wollte, als die Zusammensetzung der Flüssigkeit.

24ab Der Mathelehrer beklagt, dass die Jugendlichen nicht mehr {ausrechnen, berechnen} können, was die Wurzel aus vier ist.

24cd Der Mathelehrer beklagt, dass die Jugendlichen nicht mehr {ausrechnen, berechnen} können, als das große Einmaleins.

Curriculum Vitae

Gerrit Johannes Kentner

Adresse	Bornheimer Landwehr 84, 60385 Frankfurt am Main Tel.: 069-747 38 252 Email: gerrit@lingua.uni-frankfurt.de
Geburtsname	Gerrit Johannes Otten
Geburtsdatum	18. Dezember 1976 in Oldenburg (i.O.)
Familienstand	verheiratet, zwei Kinder

Beruflicher Werdegang

10/2009 –	wissenschaftlicher Mitarbeiter am Institut für Linguistik der Goethe-Universität Frankfurt am Main (Lehrstuhl Phonologie)
06/2006 – 09/2009	Doktorand und wissenschaftlicher Mitarbeiter (50%) am Institut für Linguistik der Universität Potsdam, angestellt im Rahmen des DFG-Schwerpunktprogramms 1234 "Sprachlautliche Kompetenz", Projekt "Prosody in Parsing"; Antragsteller: Prof. Dr. C. Féry, Prof. Dr. S. Vasishth, Dr. F. Kügler.
10/2001 – 04/2006	Studium der Allgemeinen Sprachwissenschaft (Patholinguistik), Universität Potsdam Abschluss: Diplom Allgemeine Sprachwissenschaft
09/2004 – 09/2005	Erasmusstudium (Sprachwissenschaft und Human Communication Sciences) und Forschungspraktikum an der University of Newcastle upon Tyne (UK)
10/1998 – 09/2001	Ausbildung zum Logopäden an der Schule für Logopädie der Westfälischen Wilhelms-Universität Münster Abschluss: Logopädisches Staatsexamen
09/1996 – 09/1997	Zivildienst, Annastift Hannover
06/1996	Abitur, Altes Gymnasium Oldenburg