What about those B-accents and Hat Patterns?

Form and Meaning of Contrastive Topics in English, Dutch, and German

by
Gouming Martens

A thesis submitted to the Department of Linguistics at McGill University
in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY
in
LINGUISTICS

Montreal, Canada
August 2022

其出弥远，其知弥少
Qí chū mí yuǎn, Qí zhī mí shǎo
‘The further you go, the less you know.’

老子 LAOZI

Abstract

This dissertation explores how intonation and meaning relate to one another. The following questions will be addressed: How do intonational form and meaning relate to each other? Is the mapping between intonational form and meaning compositional or should it be treated holistically? To get a better understanding of how intonation works in general, I will look into two specific intonational patterns, namely the hat pattern in German and Dutch and AB-BA patterns in English. Some have linked these intonation pattern to notions such as contrastive topics and foci where specific pitch accents are consistently associated with either topics or foci. Others take a more holistic approach and assign a specific meaning to the whole tune as such. A lot has been written about these patterns but so far there is little consensus on the exact forms and meanings. Part of the problem lies in the fact that there is no established and uniform phonological form that corresponds to these patterns: Different researchers refer to different phonological forms or fail to even mention what the form is in some cases. Another issue is that little empirical data has been provided to support a certain analysis: Most approaches linking intonation patterns to specific meanings have only informally provided individual examples to support their claims. Through a series of experiments, this study makes an attempt at exploring how intonational form and meaning are related to each other and whether intonational meaning should be considered holistic or composed of smaller meaningful units.

Résumé

Cette thèse explore les relations entre l'intonation et la signification. Les questions suivantes seront traitées: Comment la forme d'intonation et la signification sont-elles liées l'une à l'autre? La correspondance entre la forme et le sens est-elle composée ou doit-elle être traitée de manière holistique ? Pour mieux comprendre le concept d'intonation en général, je me pencherai sur deux motifs d'intonation, à savoir le motif du chapeau en allemand et en néerlandais et le motif AB-BA en anglais. Les deux motifs sont liés à des concepts tels que les sujets et les foci contrastifs, où des accents tonaux spécifiques sont systématiquement associés à des sujets ou à des foci. D'autres adoptent une approche plus holistique et attribuent une signification spécifique à tout le motif d'intonation en tant que tel. Beaucoup de littérature a été consacrée à ces motifs, mais jusqu'à présent, il n'y a guère de consensus sur les formes et les significations exactes. Une partie du problème réside dans le fait qu'il n'existe pas de forme phonologique établie et uniforme correspondant à ces motifs: Différents chercheurs font référence à des formes phonologiques différentes ou ne mentionnent même pas la forme dans certains cas. Un autre problème est que peu de données empiriques ont été fournies pour soutenir une certaine analyse : La plupart des approches liant les formes d'intonation à des significations spécifiques n'ont fourni que des exemples informels à l'appui de leurs analyses. Par le biais de plusieurs tests, cette étude tente d'explorer la manière dont la forme et la signification des intonations sont liés et si la signification des intonations doit être considérée comme un concept holistique ou composée d'unités significatives plus petites.

Acknowledgements

First and foremost, I would like to thank my thesis supervisors, Michael Wagner and Francisco Torreira. Without their assistance and dedicated involvement in every step throughout the process, this thesis would not have come into being. I would also like to thank Junko Shimoyama, who was my supervisor for the first two years of the PhD program, for her very valuable comments on this dissertation.

Also, a special thanks to my cohort members, Emily Kellison-Linn, Yeong Woo Park, Clint Parker and Jia'er Tao. I have very fond memories of our time together (especially our homework trips to a cabin the woods has stuck with me). You guys were my support and refuge, and my time as a PhD student would not have been as fun or precious without you guys.

List of Abbreviations

AM	Autosegmental-Metrical framework
CT	Contrastive Topic
EP	Early Peak
ER	Early Rise
F(oc)	Focus
IP	Intonational Phrase
LP	Late Peak
RFR	Rise-Fall-Rise
ToBI	Tones and Break Indices
ToDI	Transcription of Dutch Intonation

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The quote at the beginning of this dissertation gives a good idea of what science is like: The more you research a certain topic the more questions come to mind and the less one seems to know. This is especially the case for a topic like intonation, which only began to receive serious attention from linguists starting from the 70s. The term intonation can be defined as the use of suprasegmental phonetic features to convey 'post-lexical' or sentence-level meanings in a linguistically structured way (Ladd 2008: p.4).

Intonation is a feature of human language found amongst all spoken languages. And yet, despite this universality and omnipresence in language, we still have a rather poor understanding of how intonation systems actually work. There are still some fundamental questions that need to be answered before any substantial progress can be made. For instance, how does meaning map onto intonation patterns? In other words, what constitutes an intonational morpheme? It could be that intonational meaning is conveyed by a complete intonational tune (without internal structure), or perhaps smaller phonological units that make up a tune each carry their own semantics. In this study I will focus on how intonational form and meaning relate to each other: Does a complete intonational tune convey a certain semantics (holistic view) or do the smaller phonological units that make up an intonational tune each carry their own meaning (compositional view)?

To get a better understanding of how intonation works and more specifically how meaning maps onto intonation, I will look into two specific intonation patterns: the so-called "hat" pattern in German and Dutch[1], and Jackendoff (1972)'s BA and AB patterns in English. These two intonation patterns are said to share a similar semantics and are therefore often discussed together (Büring 1997; Constant 2014; Gast 2010; Wagner 2012). Based on this proposed semantic similarity and its importance throughout the intonational literature, I have chosen to research these two specific intonation patterns in more detail.

The hat pattern refers to an intonation pattern which starts with a rise, remains level and eventually falls, thus forming an apparent hat shape in the pitch trajectory (see Fig. 1.1). The B accent in Jackendoff (1972)'s BA and AB patterns is realized as a rise-fall-rise contour (or a rising contour depending on the author)[2] and the A accent is realized as a fall in pitch (see Fig. 1.2 and 1.3). [3]

A lot has been written about the hat pattern and AB-BA patterns, but there is little consensus on the exact forms and meanings that go with these intonation patterns. Part of the problem lies in the fact that there is no established and uniform phonological form that corresponds to these intonation patterns: Different researchers refer to different phonological forms or fail to even mention what the form is in some cases. Another issue is that little empirical data has been provided to support a certain analysis: Most approaches linking intonation patterns to specific meanings have only informally provided individual examples to support their claims. Through a series of experiments, this study makes an attempt at exploring how intonational form and meaning are related to each other and whether this relationship is best approached compositionally or holistically.

[1] Also known as the "bridge" accent (Büring 1997) or the "flat hat" contour ('t Hart and Collier 1975).
[2] With contour, a recurrent tonal pattern that aligns with the metrical structure of an utterance is meant.
[3] Although Jackendoff (1972) refers to these as pitch accents, from his description of the individual accents it becomes clear that "pitch accent" here refers not just to tonal targets associated with accented syllables, but also to tonal targets on the boundaries of certain phrases. In this study we will stick to Jackendoff (1972)'s use of the term "pitch accent" while keeping in mind that the term it sometimes includes boundary tones (if there are any).

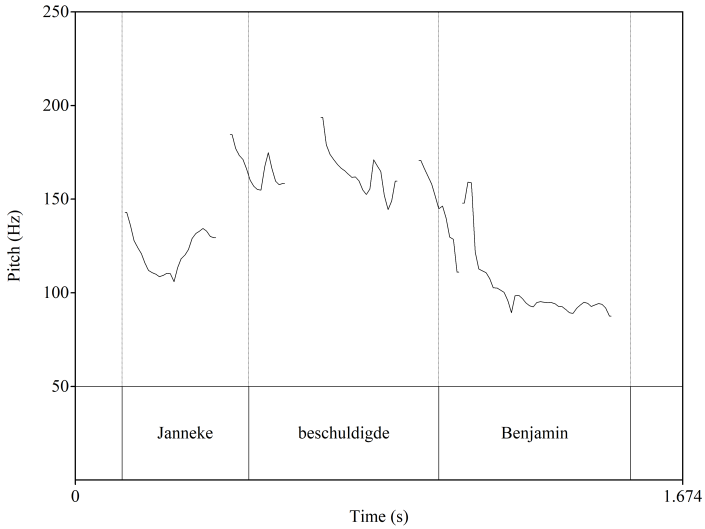


Fig. 1.1: Illustration of a hat pattern. A simple SVO sentence in Dutch (*Janneke beschuldigde Benjamin* 'Janneke blamed Benjamin') is shown in this figure, with a rising accent on *Janneke* and a falling accent on *Benjamin*. The solid black line shows the pitch track over the course of time in seconds, the dotted lines indicate word boundaries. Click here to listen.

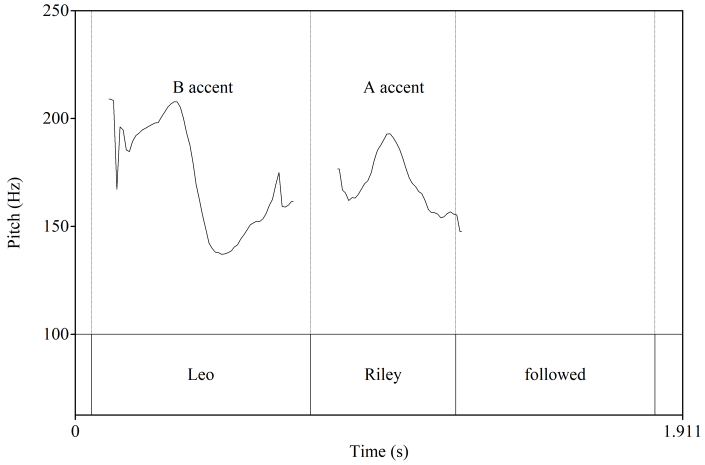


Fig. 1.2: Illustration of a BA pattern. In this figure, a topicalized sentence in English (*Leo Riley followed*) is shown. Click here to listen.

Of course there is no way to cover every aspect of intonation, but by better comprehending these two specific intonation patterns I hope to gain a better understanding of intonation in general. This dissertation carefully describes the phonetic and phonological forms of the aforementioned patterns in English, German and Dutch, and makes an attempt at unraveling the relationship between intonational form and meaning through

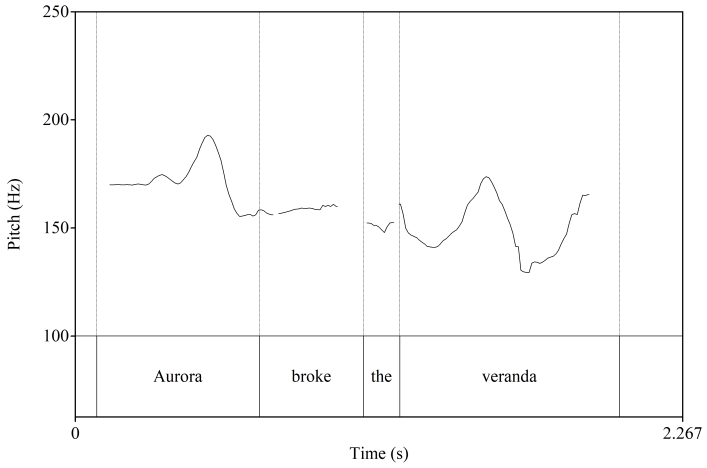


Fig. 1.3: Illustration of an AB pattern. In this figure Aurora is produced with an A accent and *veranda* with a B accent. Click here to listen.

the use of linguistic experiments: The aim is to find a good balance between the theory and empirical evidence to support (or refute) certain theories. The main research question that will be referred back to at the end of each chapter and that forms the red thread throughout this dissertation is: "How does meaning map onto intonation?".

Jackendoff (1972)'s analysis of AB-BA patterns in English will form the starting point of our narrative. In order to better understand the hat pattern in German and Dutch it is important to have an understanding of contrastive topics in English as described by Jackendoff (1972), see Example (1) below. In such sentences it is argued that there is a Contrastive Topic (in this case *Fred*) and a Focus (in this case *beans*).[4] Sentences involving a contrastive topic and focus have been closely associated with AB-BA patterns in English (Büring 1997; Constant 2014; Jackendoff 1972) and the hat pattern in German (Büring 1997; 1999; Féry 1989). In English, Jackendoff (1972) has argued that contrastive topics are marked by B accents and foci by A accents. Consequently, when contrastive topics precede foci BA patterns are expected, and when contrastive topics follow foci, AB patterns are expected. Büring (1997) and Constant (2014) follow Jackendoff (1972) in arguing that B and A accents mark for contrastive topics and foci respectively. Analogous

[4]Later, the terms *topic* and *focus* will be discussed in more detail.

to Jackendoff (1972)'s analysis for English, Büring (1997) has argued for a similar analysis of hat patterns in German: Contrastive topics are marked with rising accents and foci with falling accents. The general claim made by Jackendoff (1972), Büring (1997) and Constant (2014) in terms of how intonational form relates to meaning, is that there is a direct relation between a pitch accent and a specific meaning. From this it follows that in their view intonational meaning is compositional, in the sense that pitch accents that make up an intonational tune each carry their own meaning. Put differently, pitch accents are considered intonational morphemes.

(1) (Jackendoff 1972: 261)

A: Well, what about FRED? What did HE eat?

B: FRED ate the BEANS.

In contrast to Jackendoff (1972) and Büring (1997), Wagner (2012) takes a completely different approach to AB-BA patterns in English and hat patterns in German. Instead of relating these patterns to contrastive topics, Wagner (2012) assigns a more global meaning to these patterns, related to evoking alternative propositions. For English, Wagner (2012) claims that the BA and AB patterns are fundamentally two distinct intonation patterns that each come with their own semantics.

Both for the AB patterns (or RFR contour as Wagner calls it) in English and hat patterns in German, Wagner (2012) considers a more holistic approach by arguing that the two tunes in their entirety convey a certain meaning. The AB pattern in English supposedly conveys that alternative propositions are possibly true whereas the hat pattern (with an early peak accent) in German conveys that alternative propositions must be true. Crucially, Wagner (2012) does not assign meaning to the individual pitch accents that make up these intonation patterns, but to the whole tune itself.

The BA pattern on the other hand, is claimed to not encode any alternative-sensitive meaning, instead it is simply a result of the fact that in a declarative sentence with a falling accent, non-nuclear accents can be realized with a rise (resulting in a B-realization), while

final accented constituents are not (resulting in the A-realization). In other words, this means that the BA pattern is not really on par with the AB pattern, and in fact encodes nothing with respect to the contrastive topic status of either argument. That is not to say that AB-BA patterns and hat patterns are not compatible with sentences containing contrastive topics, however they do not specifically encode for contrastive topics prosodically.

In this study the different analyses to AB-BA patterns in English and hat patterns in German (and Dutch) will be put to the test through a series of experiments. So far no extensive experimental study has been conducted to empirically support any of the aforementioned analyses and thus there is a need for more experimental research: The aim of the current study is to provide experimental evidence to support (or reject) the claims made by Jackendoff (1972), Büring (1997) and Wagner (2012).

This first chapter lays the foundations to understanding intonation by introducing several basic notions and definitions that most researchers agree on, and assumptions that will be made throughout this study. In addition, the notion of *(contrastive) topics* as described by Jackendoff (1972) will be introduced, which, as we have seen earlier, serves as our starting point for understanding the form and meaning of AB-BA patterns in English and hat patterns in German and Dutch.

1.1 Prosody

Prosody is a generic term for the full array of acoustic features that accompany individual phonetic segments. These acoustic features mark for instance lexical stress, lexical tones and most importantly for our purposes sentence intonation. In auditory terms these prosodic features are cued by pitch (e.g. low or high) , length (e.g. long or short), and loudness of a sound (soft or loud). In acoustic terms, pitch translates to the fundamental frequency (in Hertz), length to duration (often in milliseconds), and loudness to inten-

sity/amplitude (in Decibel).[5] Since prosody does not concern itself with the individual phonetic segments but rather the properties of larger speech units consisting of these segments, prosody is characterized as being suprasegmental (i.e. on top of/orthogonal to the segments).

Various prosodic features can be applied at different linguistic levels to form prosodic patterns by themselves: On a syllable level, one can have weaker and stronger syllables (this is often achieved by changing the intensity or duration of a syllable) to make an iambic (weak-strong) or a trochaic (strong-weak) rhythmic pattern. On a word level these strong and weak syllables can be combined to make words with different stress patterns (initial stress, final stress, penultimate stress). Tonal languages are another example of how word-level prosody is employed to make lexical distinctions through pitch, e.g. 诗 *shī* [ʂɨ 55] 'poem' and 食 *shí* [ʂɨ 35] 'to eat' in Mandarin. Finally, pitch can also be employed above the word level to affect the meaning/attitude of a utterance, this is called intonation and this will be the focus of the current dissertation.

With regard to the meaning component, prosody may reflect various features of the speaker or the utterance: From para-linguistic notions like the emotional state of the speaker or the presence of irony and sarcasm to more linguistic notions like syntactic phrasing or semantic focus marking. This dissertation solely concerns itself with linguistic notions of prosody, i.e. prosody that results in a difference in meaning rather than a difference in attitude/emotion of a certain utterance.

As briefly mentioned in this section, pitch plays a crucial role in intonation. But how exactly does a pitch trajectory align with the segmental string of an utterance? As it turns out, there are certain anchor points in an utterance with which a pitch trajectory aligns. Prominent syllables and phrase edges constitute such a possible anchor points. But what does it mean for a certain syllable to be prominent?

[5]It should be noted that this list of prosodic features is not exhaustive and there are other prosodic features like voice quality and speech tempo, but the aforementioned features are sufficient for our purposes.

1.2 Prominence, Accent and Stress

In the previous section we have touched upon the idea of weak and strong syllables. In this section I would like to be more specific about how weak and strong syllables relate to notions such as *prominence*, *stress* and *accent*. When we say that a certain syllable is stronger than another syllable, we mean that that syllable is more prominent relative to other syllables. *Prominence* is a speaker's intuitive perception of strength in an utterance and it is always relative to other elements in a certain speech unit: For example, in a trochaic rhythm the initial syllable is more prominent than the following syllable, in a language with lexical stress the most prominent syllable in a word determines the stress pattern and in intonation, metrically prominent syllables in a sentence determine the metrical structure of a sentence which determines the location of pitch accents.

The terms *stress* and *accent* are often used interchangeably in the literature which has caused a great deal of confusion (Van der Hulst 2011). Bolinger (1958) and Jassem and Gibbon (1980), for instance, regard *stress* as an abstract category and *accent* as its physical manifestation. Abercrombie (1976) and Laver and John (1994), on the other hand, hold opposite views and regard *accent* as the abstract category and *stress* as the actual physical occurrence. Still others like Cutler (1984) view *stress* as a property of words and *accent* as a property of sentences.[6] To avoid any confusion I will follow Cutler (1984)'s notions of *stress* and *accent* and take *stress* to be a property on the word-level and *accent* a property above the word-level. In a way *stress* can be seen as prominence on the word level whereas *accent* is prominence on the level of intonation.

The notion of 'prominence' can be approached from three distinct perspectives: There is the sequential (or syntagmatic) perspective which refers the position of a prominent syllable relative to that of previous or following syllables (initial, final, penultimate etc.), then there is the paradigmatic perspective which concerns itself with different possible levels of prominence that can be applied to the same syllable (e.g. word-level or above the

[6]For a comprehensive review see Fox et al. (2000).

word-level) and finally there is the physical aspect according to which different phonetic cues can signal the prominence of a syllable. As we have seen before, this research follows the terminological tradition in which a difference along the paradigmatic aspect of prominence determines whether prominence is regarded as stress or accent: Stress corresponds to prominence on the lexical level and accent corresponds to prominence above the lexical level (i.e. phrases and sentences).

In terms of the syntagmatic properties of *accents*, I will follow the British tradition in making a distinction between *nuclear*, *pre-nuclear* and *post-nuclear* material (Kingdon 1958; O'Connor and Arnold 1973). According to the British tradition the nuclear accent is the last accent in an intonational phrase (IP) and is considered obligatory (i.e. each IP must have one). I use a somewhat different definition of nuclear accent and consider the nuclear accent to be the last accent in a sentence not in an IP: As we will see in this study, the domain over which nuclear accents have an influence is larger than an IP. Accents within the same domain preceding this nuclear accent are called pre-nuclear accents, and accents following this nuclear accent are called post-nuclear accents.

Turning our attention to the physical (i.e. acoustic) aspect of prominence, there are various phonetic means to realize prominence. When pitch is the main cue to an accent/stress it is called a *pitch accent*[7], when loudness is the main cue to an accent it is called a *dynamic accent* and finally when length is the main cue to an accent it is called a *quantitative accent* (Van der Hulst 2011). In this study we will only make use of the term pitch accent.

For instance, the main acoustic cue for stress in a language like Japanese is pitch and therefore Japanese is referred to as a pitch accent language, e.g. *háshì* with a HL tonal pattern means 'chopsticks' whereas *hàshí* with a LH tonal pattern means 'bridge'. Stress in languages such as Dutch and English has often been referred to as dynamic stress (Sweet 1929; Bloomfield 1994), meaning that its primary acoustic correlate is intensity. However Sluijter (1995); Sluijter and Van Heuven (1996) found that duration turns out to be the

[7]In our terminology a pitch accent on the lexical level should actually be called pitch stress, however the convention is to call these pitch accent regardless of the paradigmatic status of prominence.

most reliable acoustic correlate of stress rather than intensity, i.e. stressed syllables are longer than unstressed syllables. In terms of acoustic correlates of intonational accent, Sluijter and Van Heuven (1996) found that overall intensity is a correlate of accent rather than stress. Of course another prominent phonetic cue to intonational accent is pitch to which we will turn our attention later in this chapter.

1.3 Metrical Structure

In the previous section the notions of prominence, stress and accent were introduced, in this section we will examine how prominence is structured to create the metrical structure of a sentence, and how this metrical structure relates to stress and accent.

Intonation can be understood as involving three orthogonal dimensions: There is the 'tune' aspect or tonal aspect of intonation, then there is the phrasing aspect and finally there is the 'metrical' aspect of intonation. This study only looks into the tonal and metrical aspect of intonation. In this section we concentrate mainly on the metrical aspect of intonation. These two dimensions, i.e. tonal (intonational, consisting of tonal units/targets) and metrical structures, are independent from each other: One can have the same metrical structure with different intonational tunes, and vice versa.

The metrical dimension of intonation is concerned with determining the prominence structure of a sentence, i.e. which elements get assigned prominence and how prominent the elements are relative to each other. This mechanism determines how in a certain domain prominence is assigned and which syllables are potentially accented and thus potentially receive a pitch accent.

On a word level the prominence structure determines where lexical stress falls. In Example (2) below, prominence can either fall on the first syllable in which case it is interpreted as a noun, or prominence can fall on the last syllable in which case it is interpreted as a verb: The most prominent syllable in a word receives stress (i.e. the prominent sylla-

ble is produced relatively longer in duration and louder in intensity than other syllables in a word). Following a common practice, the metrical structure of a word or sentence is represented by a layered grid in which prominence is marked by an 'x' (Liberman and Prince 1977).

(2) a. [ob ject] *verb*
x (lexical stress)

b. [ob ject] *noun*
x (lexical stress)

As already mentioned, the metrical structure of a word/sentence can be seen as a layered grid where at each layer prominence can be assigned. Prominence can be stacked to create a more prominent syllable among other prominent syllables, this idea of stacked prominence is best illustrated with words containing a secondary stress (see Example 3). In (3) the first layer of the metrical structure marks the prominence based on the lexical stress of the individual words, the second layer marks the primary stress in a compound word by assigning an extra prominence on as already prominence-marked syllable. The syllable with two stacked prominences would receive primary stress and would be perceived as the most prominent syllable in the word whereas the other prominent syllable receives secondary stress.

(3) [A me ri can Air lines] *compound noun (company name)*
x x (lexical stress)
x (primary stress)

So far we have seen that on a word level there are several factors that determine the metrical structure of a word: there is lexical stress and primary stress assignment in compounds. The prominence structure of a word determines potential places for prominence to be assigned on a sentential level. On a sentential level *focus* is considered an important factor determining the assignment of prominence. Roughly speaking, *focus* is used to highlight a specific part of the sentence, thereby often evoking a contrast with other entities that might fill the same position (Gundel and Fretheim 2004).

In Example (4a), *contrastive focus* gives rise to a metrical structure similar to the compound word with secondary stress in Example (3): In this specific example *trains* is contrasted with *airlines*. Similarly a contrast could be made between *British* and *American* (see Example 4b) or between two compound words in which there would be an extra lexical prominence layer (see Example 4c)[8].

(4) a. *I don't like American trains but I do like...*

[A me ri can air lines] *adjective noun*
x x (lexical stress)
x (focus)

b. *I don't like British airlines but I do like...*

[A me ri can air lines] *adjective noun*
x x (lexical stress)
x (focus)

c. *I don't like British Airways but I do like...*

[A me ri can Air lines] *compound noun*
x x (lexical stress)
x (primary stress)
x (focus)

In terms of how *focus* is realized, I will assume that focus influences the metrical grid which will be aligned with a certain contour. The pitch accents specified in the contour on its turn will align with metrically prominent syllables. As discussed earlier unlike lexical pitch accents in Japanese, it is less clear how to describe these sentential pitch accents and how many there are for a given language. Most important for our purposes is to keep in mind that the type of pitch accent is orthogonal to the metrical structure of the sentence itself. Both words in (5) presumably have the same metrical structure, but focus is realized with a different pitch accent: In the question there is a low pitch accent on *airlines* whereas in the declarative sentence there is a falling pitch accent on *airlines*. There is nothing in the metrical structure itself that can predict type of pitch accent. All the metrical structure can predict is where the pitch accent will occur. Note that this is substantially different from

[8]In this example the specific airline companies are meant.

how lexical pitch accent works, for lexical pitch accents the type of pitch accent can be predicted by the phonological of a word. The separation of tune and rhythm is a crucial idea within the Autosegmental-Metrical (AM) theory which will be discussed later in this chapter.

(5) a. *Do you like...*

[A me ri can	Air lines]		*compound noun*
x	x		(lexical stress)
x			(primary stress)
x			(focus)
L		H%	

b. *I don't like British Airways but I do like...*

[A me ri can	Air lines]		*compound noun*
x	x	(lexical stress)	
x			(primary stress)
x			(focus)
HL		L%	

So far we have observed cases with only one focus in the sentence. It is of course possible to have multiple foci in a sentence (see Example 6). The expectation in such cases is that each focused constituent is realized with a pitch accent. Recall from the previous section that this study makes a distinction between nuclear and non-nuclear accents, where the last accent in an utterance constitutes the nuclear accent: *Fred* in Example (6) for instance, would be considered pre-nuclear. As we will see throughout this study, the notions nuclear, pre-nuclear and post-nuclear will play a significant role when it comes to the actual accent realizations of a sentence. In this study we are mainly interested in sentences with multiple accented constituents since these would be the realizations in which AB-BA patterns and hat patterns can be produced in their full form.

(6) *Who ate what?*

Fred	ate the	beans.	*double focus*
x	x	x	(lexical stress)
x		x	(focus)
preNuc		Nuc	

Although focus has a certain influence on the actual accent realization of a sentence it does not definitively determine how the sentence is realized in terms of pitch accents. For instance, cases have been reported of non-focused material receiving a pitch accent and vice versa.[9] In this study we will see that the metrical structure induced by focus does not always match up with the pitch accent realizations expected by this metrical structure.[10]

Having discussed the metrical structure of an utterance which determines the potential anchor points of a pitch trajectory, one can now focus on the tonal aspect of intonation.

1.4 Tone and Intonation

Tones are phonological labels used to categorize pitch trajectories that are perceived to be similar. In tonal languages like Mandarin or Vietnamese there is a specific number of tonal categories that are easily perceived by native speakers since they are crucial in marking lexical distinctions: A change in tone can bring about a change in meaning, e.g. 诗 *shī* [ʂi 55] 'poem' and 食 *shí* [ʂi 35] 'to eat' in Mandarin. In a similar way to tone, intonation refers to recurring pitch patterns on a phrasal level, each of which is used with a set of relatively consistent meanings (Cruttenden et al. 1997). However, there are few intonation patterns that have such a relatively consistent meaning that everyone agrees on: It is far less clear what it means to be distinctive on a phrasal-level compared to on a word-level. For instance, often certain meanings associated with a specific intonation pattern are not exclusively indicated by intonation but might be encoded in the syntax as well, e.g. a question may be indicated by a rising intonation at the end of a sentence or by subject-auxiliary inversion in English. Moreover, there is no agreement on what constitutes an intonational morpheme and what kind of meanings can be ascribed to these intonational morphemes: The same intonation pattern for instance can supposedly express various

[9] See Chapter 7 of Ladd et al. (2009) for a detailed discussion on this matter.

[10] This discrepancy between metrical structure and accent realization seems to be related to the notion of nuclear, pre-nuclear and post-nuclear accents as we will see in later chapters.

(unrelated) meanings, e.g. a final rising intonation often marks a question at the end of a sentence, but it can also mark a continuation at the end of a phrase. Even labels like question intonation and continuation intonation are problematic since questions actually take all sorts of tunes and there is no established "question" tune.[11]

The current study focuses specifically on the issue of how intonational meaning is encoded by a certain intonation pattern, in other words what constitutes an intonational morpheme.

1.4.1 Compositionality

Some accounts regard intonation as 'holistic' in the sense that entire tune is meaningful, i.e. the tune cannot be decomposed into smaller meaningful units (see Delattre (1966) for French and Liberman and Sag (1974) for English). In this specific study a tune refers to a pitch pattern over an entire sentence (possibly spanning multiple successive IPs). According to these accounts there is a correspondence between tunes and meanings. For instance, there is the question tune, the exclamation tune, the continuation tune and many more meaningful tunes. Wagner (2012)'s analysis of the hat pattern and the AB pattern can be considered a 'holistic' approach: Intonational meaning is ascribed to the whole tune and it is the complete tune itself which is considered to be the smallest meaningful unit (i.e. intonational morpheme).

Then there is the so-called partly 'holistic' view according to which tunes have internal morphological structure, i.e. they can be decomposed into smaller meaningful units and these meaningful units can be combined to create tunes. Bolinger (1958) for instance, recognizes three accents in English, each with their own semantics: A accents, B accents and C accents. The A accent describes a falling pitch gesture and conveys assertiveness, it is used with items that are separately important, contrastive, and/or new to the discourse (Bolinger 1958; Bolinger and Bolinger 1986). The B accent describes a rising pitch ges-

[11] For an experimental study to this specific issue I would refer the reader to Torreira and Valtersson (2015).

ture and it conveys something like 'connectedness' and 'incompleteness'. The C accent is a low and flat contour and it is described as an 'anti-Accent A', it conveys something like 'anti-assertiveness' which Bolinger (1958) describes a lack of spirit or motivation. Other partly 'holistic' approaches are the British school of intonation research (Crystal 1969; O'Connor and Arnold 1973) and the Institute for Perception Research (IPO) Intonation Grammar which is based on the British school of thought ('t Hart et al. 1990). According to O'Connor and Arnold (1973) there are certain "tone groups" with which meaning is associated and these tone groups are composed minimally of a nucleus and optionally of a pre-nuclear and post-nuclear part. The approaches taken by Jackendoff (1972), Büring (1997) and Constant (2014) to the AB-BA pattern and hat pattern can be considered partly holistic: Meaning is assigned to specific pitch accents[12] that make up a tune and thus according to these approaches accents form the smallest meaningful intonational unit. According to Jackendoff (1972), Büring (1997) and Constant (2014), the A and B accents have consistent meanings across the AB-BA patterns. The global meaning of AB-BA patterns is compositionally derived from the meanings assigned to the individual pitch accents.

In a fully compositional analysis an intonation pattern is composed of individual tones each with specific meanings (every tonal unit at the phonological level is a morpheme). Pierrehumbert and Hirschberg (1990) were the first to propose a compositional analysis of intonation within the autosegmental-metrical (AM) framework.[13] In this strictly compositional approach every tone and bitonal combination is assigned a meaning. A high tone aligned with the accented syllable (H*) for instance indicates that the accented referent should be treated as 'new' in the discourse and high tones aligned with the boundary of a phrase (H- or H%) convey that the phrase is to be interpreted in relation to a later phrase of the same type; a low boundary tone (L%) on the other hand indicates that the phrase is to be interpreted separately from any other sentence. Another example of a fully com-

[12] As discussed earlier boundary tones may or may not be included.
[13] More about the AM theory later in this chapter.

positional analysis is Gussenhoven (2016a)'s approach to intonation, like Pierrehumbert (1980) each tone and bitonal combination is assigned a meaning.

For the sake of simplicity and length of this thesis, the current study solely concerns itself with holistic and partly holistic approaches to intonation where meaning resides in intonational phrases/groups or combinations of these rather than individual pitch accents and boundary tones, with Wagner (2012)'s account leaning more towards a holistic approach and Jackendoff (1972), Büring (1997) and Constant (2014) leaning more towards a compositional approach.

So far, the meanings ascribed to these contours are rather vague and lack a consistent formal representation. This is an issue found throughout the intonation literature: Phonetically oriented literature tends to capture a certain intonation pattern fairly well in terms of the actual realization, but they lack a rigid understanding of the meanings of such patterns. And vice versa, semantically/pragmatically oriented literature do a good job of making formal semantic analyses, but they often fail to elaborate in more phonetic/phonological detail which intonation pattern is actually investigated. The danger here lies in the fact that one might come up with one formal semantic analysis which is based on a variety of seemingly similar but phonologically different patterns: If all these different phonological forms have their own semantics then trying to come up with one overarching semantic analysis will not make much sense. The other way around is as problematic, i.e. there are different phonological forms but all are referred to with the same label. This is exactly what is happening in case of the hat pattern: Researchers formulate semantic analyses for 'the' hat pattern, often without actually mentioning its phonological form. Therefore it might very well be the case that one researcher has a completely different hat pattern in mind than his/her colleagues although both refer to it as 'the' hat pattern. The main goal of this dissertation is to tackle this problem by clearly defining which patterns are being investigated before analyzing the meanings of these patterns.

1.5 Autosegmental-Metrical Theory (AM theory)

Currently the most widespread phonological framework for representing intonation is termed the 'autosegmental-metrical (AM)' framework. This framework started with the work of Pierrehumbert (1980), who was herself inspired by the work of Bruce (1977) on Swedish and Liberman (1978) on English. The term autosegmental-metrical was coined by Ladd (2008) and it refers to the separation of tones and prominence on different planes of the phonological representation. The metrical aspect of the autosegmental-metrical theory refers to the the division of utterances into phrases and the assignment of relative prominence to elements within the phrase (phrasing and highlighting), which was first proposed by Liberman and Prince (1977). The autosegmental aspect refers to the association of the tones with the metrical structure. Pierrehumbert (1980) were the first to propose a compositional analysis of intonation within the autosegmental-metrical (AM) framework (Féry 2017). As discussed before, in this strictly compositional approach every tone and combination of tones is assigned a meaning. In Pierrehumbert (1980)'s framework intonation contours are decomposed into relatively high and relatively low pitch levels: High (H) and low (L) tones. 'Relatively low/high' means low/high relative to the local phrasal pitch range. A crucial aspect of Pierrehumbert (1980)'s analysis is the distinction between pitch accents (H*, L*, L+H* and L*+H), phrase accents (H- and L-), and boundary tones (H% and L%). Building on Pierrehumbert (1980)'s framework, the Tones and Break Indices (ToBI) system was created.

1.5.1 Tones and Break Indices (ToBI)

The most widely used annotation system for representing intonation is the ToBI system. ToBI started out as an annotation system used to annotate American English (also known as MAE_ToBI) (Jun 2007), but soon it developed into a general theoretical framework for analyzing intonation systems of other languages (e.g. Grice et al. (1996) for Ger-

man (GToBI); Venditti (1997) for Japanese; Gussenhoven et al. (2005) for Transcription of Dutch Intonation (ToDI)). Because of the wide-spread use of ToBI across different languages within the field of intonation research, the current study will follow this practice and use the ToBI annotation system for representing intonational tunes. This is a more or less arbitrary choice and another convention could have been used as well. However, given that most studies referred to in this thesis make use of the ToBI annotation system and given that ToBI is adequate enough to represent the contours dealt with in this study, it makes sense to follow this convention.

ToBI developed from the theoretical framework set out by Pierrehumbert (1980) and therefore it shares many similarities with Pierrehumbert (1980)'s framework. The ToBI system can be characterized by the following five features: First, since ToBI is at its core based on the AM theory, the intonation contours and metrical structure are represented on different tiers of the phonological representation. The intonation contour is represented linearly by an autosegmental string of tones, whereas the metrical hierarchy of intonational phrases and lower-level prosodic groupings are represented by a numerical break index value which represents the perceived degree of disjuncture between any two words (Jun 2007). The current study will mainly make use of the tonal tier when describing and analyzing intonation contours, as is the convention in the field.

Second, the fundamental building blocks of intonation contours are relatively high and relatively low pitch levels, i.e. high (H) versus low (L) tones: Relatively low means low relative to the local phrasal pitch range and idem dito for high tones. Crucially, the ToBI system is based on pitch levels (tone targets) rather than pitch changes (dynamic tones), this is in contrast to for instance Bolinger and Bolinger (1986) and 't Hart et al. (1990) who make use of pitch changes as fundamental building blocks of intonation rather than pitch levels. However, by combining these two atomic tones dynamic tones (or bitonal combinations), i.e. rising and falling pitch trajectories, can be formed in ToBI.

Third, the local pitch range is determined by a variety of factors, such as declination and para-linguistic factors (i.e. happy speech might have a broader range than bored speech): Because of this it can happen that a certain H tone in one part of the intonation contour is lower (in an absolute sense) than a L tone somewhere else in the same utterance (Jun 2007).

Fourth, tones are either edge tones or pitch accents: a tone aligned with an accented syllable is a pitch accent whereas an edge tone is aligned to the segments at the relevant phrase boundary. The f0 between adjacent tonal targets is obtained by linear interpolation.

The fifth feature of ToBI mentioned by Jun (2007) is somewhat controversial and not shared by all ToBI accounts. The basic idea is that edge tones can be divided into two categories based on the level of intonational phrasing they align with: Edge tones that align with intonational phrases are called *boundary tones*, these phrases often coincide with the end of a sentence. Edge tones that align with intermediate phrases are called *phrase accents*, these phrases often coincide with syntactic phrases. However as mentioned before not all accounts of ToBI recognize two edge tones, Gussenhoven (2016a)'s model for American English and Dutch for instance only has boundary tones but no phrase accents. More importantly, there is very limited evidence that two levels of tonal phrasing are needed rather than one or more than two (Wightman et al. 1992). It is however beyond the scope of this dissertation to determine whether or not intermediate phrasal tones exist or not. In this thesis I follow Gussenhoven (2016a) by assuming there is only one group of edge tones, namely boundary tones.

More generally, this study follows Gussenhoven (2016a)'s ToBI system to represent intonation patterns. One should however keep in mind that for the purposes of this study which specific phonological framework is followed is less important. Gussenhoven (2016a)'s ToBI was chosen based on its simplicity and intuitiveness. Notice that by following Gussenhoven (2016a) there is no intermediate phrasal tone in the phonological representation and there are no so-called leading tones. Leading tones are tones that de-

scribe the pitch targets preceding the accented syllable. Although this study mainly follows Gussenhoven (2016a)'s phonological framework, for certain contours I will diverge from this framework and assume that there are leading tones.

1.6 Semantics of AB-BA and Hat Patterns

Having introduced the phonetic/phonological terminology and notions to intonation, one can now focus on the semantic part of intonation. Since this study only examines AB-BA patterns in English and hat patterns in German and Dutch, this section will limit itself to the semantic notions and analyses associated with these intonation patterns. As mentioned at the beginning of this chapter, our narrative starts with Jackendoff (1972)'s analysis of the AB-BA pattern. In this analysis the AB and BA patterns are associated with notions like contrastive topic and focus.

Before turning our attention to the semantics of a focused element we have to get a basic understanding of the semantics of questions and answers. We will therefore start by asking ourselves: "What is the meaning of a question?". Of course there is not enough space in this dissertation to fully answer this question, instead we will follow Rooth (1985)'s alternative semantics framework and focus on those parts that are most important for understanding the following chapters.

The semantics of questions as we know it today has only been seriously studied since the second half of the 20th century, marked by the ideas of Hamblin (1973). Hamblin (1973) formulated the idea that "questions set up a choice-situation between a set of propositions, namely those propositions that count as answers to it". Put differently, this means that a question denotes the set of all propositions, in a world w, that correspond to a possible answer to the question. But what exactly constitutes a "possible answer" to a question? And how does the notion of "possible" answers relate to notions of focus and contrastive topic?

1.6.1 Focus

From a functional point of view *focus* is roughly used to call attention to a specific part of the sentence, thereby often evoking a contrast with other entities that might fill the same position (Gundel and Fretheim 2004). [14] The question/answer paradigm has been used as a diagnostic for what counts as focused information. A clear example of focus is in the case of wh-questions: The wh-word invokes a focused constituent in the answer. If the subject is the one that is being questioned, then the subject in the answer is focused (see 7). The same goes for when the object is being questioned (see 8) or when both constituents are questioned (see 9).

(7) A: Who saw Anna? (Subject focus)

B: JOHN saw Anna.

(8) A: Who did John see? (Object focus)

B: John saw ANNA.

(9) A: Who saw whom? (Double focus)

B: JOHN saw ANNA.

Although there are several formal approaches to focus, we will follow the alternative semantics approach to focus pioneered by Rooth (1985). In this approach each focused constituent α has both an ordinary denotation $[\![\alpha]\!]_o$ and a focus denotation $[\![\alpha]\!]_f$. The ordinary denotation of a sentence is simply whatever denotation it would have in a non-alternative-based system while its focus denotation can be thought of as the set containing all ordinary denotations one could get by substituting the focused constituent for another expression of the same semantic type (Rooth 1985). Let us return then to the issue raised in the previous section of what would be considered a "possible answer" to a specific question. Following Rooth (1985)'s focus semantics, a "possible answer" to a question can be seen as any element in the focus denotation of a certain sentence, i.e. the set

[14] Although the current study solely focuses on prosodic cues to focus it should be noted that focus can also be marked through focus particles, e.g. *only* or *even* in English.

containing all ordinary denotations one gets by substituting the focused constituent for another expression of the same semantic type.

1.6.2 Contrastive Topics

Related to focus is the notion of (contrastive) topic, generally speaking the topic of a sentence is what is being talked about. Jackendoff (1972) uses a slightly altered question/answer paradigm to determine the topic and focus of a sentence (see Example 11). This paradigm consists of two consecutive questions: The 'what about'-question establishes a topic (in this case *Fred*) and the second wh-question determines the focus as we have seen in the previous section.

The constituent that is construed as the focus is realized with an A accent and the one as a topic is realized with a B accent.[15] Like Rooth (1985), Jackendoff (1972) represents focus by substituting a variable for the constituent encoding focus.

While according to Jackendoff (1972), there is only one constituent serving as primary focus and carrying an A accent, he observes that a sentence can include multiple foci. Non-primary foci are semantically represented by variable substitution, just like simple foci. A sentence with two foci introduces a presupposition defined by abstracting over a tuple of constituents (Jackendoff 1972), see the derivation in (10).

(10) λ(x,y). x ate y

The presupposition of such a sentence with multiple foci is that there is some tuple $< x, y >$ such that the proposition derived by filling in those values is true. According to Jackendoff (1972) these variables are filled in an inherently asymmetric way: The value for one variable is freely chosen (the 'independent variable'), while the other variable is filled in second (the 'dependent variable'). The dependent variable corresponds to the primary

[15] By now it should have become clear that it is more correctly to speak of the topic being aligned with a pitch accent that is part of a B accent.

focus, while the way the independent focus is interpreted captures what it means to be a 'topic'.

According to Jackendoff (1972) topics are marked by B accents (Rise-Fall-Rise contours)[16] and foci by A accents (falling contours), this is illustrated in Example (11) where one would expect a BA accent sequence. If the focus is on the first DP and contrastive topic on the second DP, the pitch accents are expected to be switched around and one would expect to find an AB pattern, as can be seen in Example (12).

Jackendoff (1972) technically implements this distinction between dependent and independent focus by assuming that syntactic F-markers are diacritically marked for whether they carry an A or a B accent, and the implicit assumption is that there can be only one constituent marked to carry the A accent. Jackendoff (1972) links the discourse function of independent foci to the notion of topic, while the dependent focus provides the sentence focus. It should become clear that Jackendoff (1972) considers an pitch accent to be an intonational morpheme, i.e. there is a direct mapping between a pitch accent and a meaning.

(11) (Jackendoff 1972: 261)

A: Well, what about FRED? What did HE eat?

B: FRED ate the BEANS.
B A

(12) (Jackendoff 1972: 261)

A: Well, what about THE BEANS? Who ate THEM?

B: FRED ate the BEANS.
A B

1.6.3 Büring (1997)

Building on Jackendoff (1972)'s topic-focus analysis of AB-BA patterns in English, Büring (1997) implements a slightly different topic-focus analysis to B and A accents in English

[16] In the next chapter a more elaborate phonological description will be provided.

and applies this same analysis to the hat pattern in German. Sentences with a topic-focus structure in German are produced with a hat pattern according to Büring (1997). The hat pattern (or the bridge pattern as he calls it) is taken to be compositional, meaning that it is composed of several morphological units. In this case the hat pattern consists of two phonological units: the initial rise and the final fall.[17] Each phonological element has its own pragmatic/semantic function, the rise constitutes the sentence internal topic (S-topic or contrastive topic in our terminology) and the fall constitutes the focus of the sentence. Whichever element in the sentence receives focus will be part of the set of propositions that are considered well-formed alternatives, this Büring calls the *focus value*: Following Rooth (1985)'s concept of alternative semantics, this can be formalized by deriving a second semantic value from a sentence S, which is called the focus value of sentence S or $[[S]]^f$ for short. If the Focus is placed on the object NP, $[[S]]^f$ is the set of propositions we get by sticking in alternatives for the focus. In Example (13) for instance, *apples* is focused and the focus value in this specific sentence could be a set like this: {*The chefs ate apples, the chefs ate pears, the chefs ate pies...etc.*}

(13) A: What did the chefs eat?
B: The chefs ate APPLES.

S-topics are similar to focus in the sense that they both induce alternatives. These alternatives however, do not have any impact on the focus value, instead, it is a set containing different focus values (Topic Value), i.e. a set of sets of propositions. The topic value is a set inside of which the focus values are nested, the topic value for an answer like in (14) could be: { {*The chefs ate apples, the chefs ate pears, the chefs ate pies...etc.*} {*The waiters ate apples, the waiters ate pears, the waiters ate pies...etc.*} }. It should become clear that the focus value is nested within the topic value and by changing the subject, the topic of the sentence changes.

[17] A more detailed description of the phonology of the hat pattern will be provided in Chapter 4 and 5.

(14) A: What did the chefs eat?

B: The WAITERS ate PEARS.

Different from other analyses on contrastive topics, according to Büring (1997), S-Topics carry a certain implicature, namely a disputability implicature. This implicature is formulated as: "Given a sentence A, containing an S-Topic, there is an element Q in $[[A]]^t$ such that Q is still under consideration after uttering A". Formulated differently this means that there is a question in the set of questions denoted by $[[A]]^t$ which is still disputable. A question is said to be disputable if there are informative but non-absurd answers to it. The question in (14) is an example of such a disputable question in the context of the answer.

Constant (2014)'s account of contrastive topics in English is very similar to Büring (1997)'s account in that both argue for a separate CT-operator and a Focus-operator. Constant (2012)'s analysis however assumes a more compositional approach to deriving the meaning associated with contrastive topics. The contrastive topic starts out as a simple focus and thus at the very start of the derivation, CT-constructions are simply cases of double focus. Through a process of CT-raising induced by a CT-operator, a focused constituent gets interpreted as a contrastive topic. The details of this derivation are not central for the purpose of this study, what is most important is that both Büring (1997) and Constant (2014) argue for a separate CT- and Focus-operator.[18]

Moreover, crucially, all three accounts consider B and A accents to mark for contrastive topics and foci respectively, and thus pitch accents are considered to be intonational morphemes.

[18]In this study, Büring (1997) and Constant (2014)'s accounts of English AB-BA patterns are not discussed separately since they make similar predictions for the stimuli found in this study.

1.6.4 Alternative Approach

In clear contrast to Jackendoff (1972), Büring (1997) and Constant (2014), Wagner (2008; 2012) contests the idea that the accents that make up AB-BA patterns and hat patterns consistently mark contrastive topics and foci. Instead Wagner (2012) argues for a more holistic approach to these two intonational patterns whose meanings although compatible with sentences containing contrastive topics, do not specifically encode for contrastive topics.

The AB pattern in English supposedly conveys that alternative propositions are possibly true whereas the hat pattern (with an early peak) in German conveys that alternative propositions must be true. Crucially, Wagner (2012) does not assign meaning to the individual pitch accents that make up these intonation patterns, but to the whole tune itself.

The BA pattern, on the other hand, is claimed to not encode any alternative-sensitive meaning, instead it is simply a result of the fact that in a declarative sentence with a falling accent, non-nuclear accents are realized with a rise (resulting in a B-realization), while final accented constituents are not (resulting in the A realization). In other words, this means that the BA pattern is not really on par with the AB pattern, and in fact encodes nothing with respect to the contrastive topic status of either argument.

1.7 Research Questions

Based on the issues and disagreements raised in this chapter, there are two main questions that this dissertations will address. First and foremost, how do intonational form and function relate to each other? Is the mapping between intonational form and meaning compositional or should it be treated holistically? In the previous sections we have seen some studies assign meaning to a whole tune (Delattre 1966; Liberman and Sag 1974; Wagner 2012) while other accounts decompose a tune into smaller units and assign meaning to specific accents (Bolinger 1958; Büring 1997; Constant 2014; Jackendoff

1972). In this study I will specifically compare Jackendoff (1972) and Büring (1997)'s compositional analyses of AB-BA patterns in English and hat patterns in German to Wagner (2012)'s more holistic approach to these two intonational patterns.

Secondly, how does the metrical structure of a sentence align with these intonational tunes? Depending on whether one takes a more holistic approach to intonation or a more compositional approach, different behaviors are expected with respect to how contours are realized given a certain metrical structure. Based on the holistic view reduction effects are expected when there is not enough metrical material for the full fledged tune to be produced (think of deletion of an accent because there is only one prominent syllable in the metrical structure of the sentence), despite such a reduction in tune the meaning should remain the same. According to the compositional view on the other hand, a deletion of an accent should lead to a difference in meaning.

To answer these questions a series of experiments was performed which aimed to test what effect (if any) focus, contrastive topics and the existence of alternative propositions (in English, German and Dutch) have on intonation.

1.8 Chapter Outlook

This chapter has laid the foundations for understanding the following chapters by introducing and defining the relevant terminology. The rest of this dissertation is split into two parts: Part I consists of chapters 2 and 3, and focuses on the AB-BA patterns in English. Part II consists of chapters 4-7 and focuses on the hat pattern in German and Dutch. Chapter 2 introduces the accents that are relevant for the topic-focus constructions in English focusing on the B and A accents. Establishing the phonological form will be crucial to getting a better understanding of the semantics of these different phonological forms. This chapter also tests the claims made by Jackendoff (1972), Büring (1997), who argue that B accents mark topics and A accents mark focus. The results show that BA patterns do not

in fact encode which constituent serves as the contrastive topic or focus, furthermore the AB pattern is rarely produced in the prototypical contexts where Jackendoff (1972) would expect such patterns. Chapter 3 discusses a different analysis to the AB-BA patterns in English and tests Wagner (2012)'s hypothesis that the Rise-Fall-Rise pattern (i.e. the AB pattern) presupposes that an alternative proposition is possibly true. The results from Chapter 3 tell us that contrary to Wagner (2012)'s predictions, the Rise-Fall-Rise pattern is not likely to be used in contexts in which an alternative is possibly true, again it is rarely used across the board.

The second part of the dissertation focuses on the hat pattern in German and Dutch. Chapter 4 introduces the accents that are relevant for the hat pattern in German and Dutch and explores the phonetic/phonological differences between two different falling accents in Dutch and German in more detail, focusing on the tonal reduction aspect. The distinction between the early and late peak will turn out to be important in later chapters. Chapter 5 tests the claims made by Büring (1997) who adopts Jackendoff (1972)'s analysis for English to German. The general idea is that in topic-focus sentences rising accents mark topics and falling accents mark foci. Unlike the AB-BA patterns in English the hat pattern in German and Dutch do seem to correlate with contrastive topic sentences, however like for the English data there was no evidence that specific accents encode for contrastive topics and focus. Chapter 6 tests an alternative approach to the hat pattern taken by Wagner (2012). According to Wagner (2012) the hat pattern with an early peak presupposes that an alternative proposition is true. The results partially supported this claim in the sense that early peak hat patterns were less acceptable in contexts without alternative propositions. Chapter 7 is similar to chapter 6 except that rather than a production experiment, a perception experiment was performed. Similarly to chapter 6, chapter 7 tested Wagner (2012)'s analysis of the hat pattern, in addition it tested Büring (1997)'s disputability implicature according to which the hat pattern presupposes that there is still an unresolved ('disputable') question. The results from chapter 6 were replicated in perception in chap-

ter 7: Early peak accents as well as rising accents were less acceptable in contexts without alternative propositions. The final chapter concludes the dissertation and discusses the results in more detail.

Part I: AB-BA Patterns

Chapter 2

Topic-Focus Constructions in English

This chapter explores sentences consisting of a contrastive topic and focus in North American English, and the intonation patterns that go with such sentences. Our starting point will be Jackendoff (1972)'s analysis of B and A accents which was briefly introduced in the previous chapter but will be explained in more detail in this chapter. The basic idea behind Jackendoff (1972)'s analysis is that contrastive topics are marked by a B accent and foci are marked by an A accent. An online production experiment was performed to test this claim. Given this analysis it is expected that constituents that constitute a contrastive topic are produced with a B accent and constituents that constitute a focus are produced with an A accent. This chapter outlines the basic claims Jackendoff (1972) made about contrastive topics in English, and work building on it, especially Büring (1997), and evaluates the validity of these claims. These accounts are compared with a more simple account in which prosody does not in fact phonologically distinguish contrastive topics and contrastive foci, rather, non-final foci can optionally be realized in a way that corresponds to what other accounts have argued are contrastive topic realizations.

2.1 The AB-BA Hypothesis

Sentences can include multiple focused constituents that contrast with alternative substitutions. Starting with Jackendoff (1972), many current accounts of contrast distinguish different types of contrastive constituents, some whose functions seem related to the notion of topicality, and which are today often called contrastive topics, and some whose functions are often characterized as sentence focus. These different kinds of contrast have been linked to different phonological realizations.

2.1.1 Jackendoff 1972

According to Jackendoff (1972), every utterance has exactly one constituent that serves as its primary focus. Which constituent serves as the sentence focus can be manipulated by using wh-questions in the context (see Examples 15 and 16).

(15) *Object Focus*

A: What did Fred eat?

B: Fred ate the BEANS.

(16) *Subject Focus*

A: Who ate the beans?

B: FRED ate the beans.

The constituent that is construed as the primary focus is realized with an A accent (a falling gesture on the accented syllable). Jackendoff (1972) makes additional assumptions about phonology that makes non-focused material less prominent than focused material, and assures that the focused constituent contains the nuclear accent. On the semantic side, focus is represented by substituting a variable for the constituent encoding primary focus, see the derivation in (17).

(17) Object focus: λx. Fred ate x

Subject focus: λx. x ate the beans

For Jackendoff (1972), a sentence with prosodic focus on a constituent x introduces the presupposition that there is some true proposition for one substitution of x, and asserts the regular meaning of the expression derived by filling in the actual value for x.[1]

While according to Jackendoff (1972), there is only one constituent serving as primary focus and carrying an A accent, he observes that a sentence can include multiple foci. In their phonological realization, multiple foci require what would today be called multiple intonational phrases. Jackendoff (1972) refers to those as "pitch accents", but his discussion makes it clear that the pitch accents are followed by boundary tones. In this study we will stick to this use of the term "pitch accent" since it would be too laborious to keep writing "pitch accent and boundary tones", but one should keep in mind that the term "pitch accent" sometimes includes boundary tones (if there are any). On the semantic side, non-primary foci are semantically represented by variable substitution, just like simple foci. A sentence with two foci introduces a presupposition defined by abstracting over a tuple of constituents (Jackendoff 1972), see the derivation in (18).

(18) λ(x,y). x ate y

The presupposition of such a sentence with multiple foci is that there is some tuple $< x, y >$ such that the proposition derived by filling in those values is true. However, according to Jackendoff (1972) these variables are filled in an inherently asymmetrical way: The value for one variable is freely chosen (the 'independent variable'), while the other variable is filled in second (the 'dependent variable'). The dependent variable corresponds to the primary focus, while the independent varialble corresponds to a 'topic'.

Which variable is the dependent variable and hence the primary focus and which is the independent variable is audible by their phonological realization: The dependent and

[1]The regular meaning of the expression is a proposition in the case of an assertion, but Jackendoff (1972) also discusses focus in polar questions, which would denote a set of propositions in many current theories.

primary focus is realized by an A accent, the independent focus by a B accent (i.e. a falling-rising contour). The crucial distinction between A and B accent is that the A accent is followed by a falling boundary, while B accent is followed by a rising boundary.

In principle, the following type of question leaves it open which focus is the independent focus, and which one is the dependent focus (Example 19).

(19) A: Who ate what?

B: Fred ate the beans.

The response should in principle be compatible with a BA and an AB pattern.[2] Jackendoff (1972) elaborates, however, that we can settle which constituent is the primary, dependent focus (and A accented) and which is the independent focus (and hence B-accented) by providing a more complex context in which an additional single-wh question determines the primary focus of the response (see 20 and 21)

(20) (Jackendoff 1972: 261)

Speaker A in the discourse is asking questions of the form Who ate what *and Speaker B is answering. For the first intonation pattern, A is asking person by person:*

A: Well, what about FRED? What did HE eat?

B: FRED ate the BEANS.
B A

(21) (Jackendoff 1972: 261)

For the second pattern, A is asking by foods:

A: Well, what about THE BEANS? Who ate THEM?

B: FRED ate the BEANS.
A B

Jackendoff (1972) technically implements this distinction between dependent and independent focus by assuming that syntactic F-markers are diacritically marked for whether they carry an A or a B accent, and the implicit assumption is that there can be only

[2] although it's possible that the BA pattern is maybe the default choice in such cases for Jackendoff (1972)

one constituent being marked to carry the A accent. Jackendoff (1972) links the discourse function of independent foci to the notion of topic, while the dependent focus provides the sentence focus. In today's terminology, Jackendoff (1972)'s independent foci are usually called contrastive topics, and dependent foci are simply called foci. Crucially, both types of foci are pitch accented, and contribute to the focus presupposition. Jackendoff (1972)'s theory predicts that in contexts like that in (20), responses should be realized with a BA pattern, and in contexts like (21), responses should be realized with an AB pattern.

2.1.2 Büring 1997

Büring (1997) provides an arguably more insightful semantics for the distinction between contrastive topics and foci. Like Jackendoff (1972), Büring (1997) assumes two diacritically distinguished types of focus markers in syntax, F (for focus) and CT (for contrastive topic). A two-step process derives a so-called topic-semantic value for every sentence, i.e. a set of questions. First, the focus-marked constituent is substituted by all contextually relevant alternatives, thus creating a set of expressions, i.e. the so-called focus value. Then the topic-semantic value is created by substituting the CT-marked constituent with particular values to create a set of sets of propositions. In other words the topic value of a sentence can be seen as a set containing different focus values which itself is a set of expressions. The presence of a CT features triggers an implicature based on the topic-semantic value, namely that there must be at least one element (a set of propositions, i.e., a question) that is still disputable in the context.

On the phonological side, constituents marked with CT are realized in English with a B-accent, and constituents marked with F with an A accent. Note that there is nothing in Jackendoff's or in Büring's theory that would force a speaker to use two intonational accents in (20) and (21). They could in principle simply use one focus, and thus only prosodically signal the relation to the antecedent provided by the immediately preceding question. We might hence expect that speakers will actually realize the responses in (20)

and (21) instead as illustrated in examples (22) and (23) below. Crucially these productions would follow the metrical structure of the sentence which is determined by the immediate wh-question itself.

(22) B: FRED ate the beans.
A

(23) B: Fred ate the BEANS.
A

The crucial claims of these two accounts about A vs. B accents applies to cases in which both constituents are accented: In such cases, prosody should disambiguate which accent signals a focus (and distinguishes the focus alternatives within each question in the topic semantic value) and which serves as a contrastive topic (and distinguishes the questions from each other). Büring (1997)'s account predicts, just like Jackendoff (1972)'s account, that in contexts like that in (20), responses should be realized with a BA pattern, and in contexts like (21), responses should be realized with an AB pattern.[3]

2.2 Double Focus Accounts

Wagner (2008; 2012) contests the idea that BA and AB patterns are equivalent interpretively. According to Wagner's analysis, Jackendoff (1972)'s and Büring (1997)'s AB patterns are actually instances of utterances pronounced with a sentence-level tune, the so-called Rise-Fall-Rise pattern (RFR). This pattern, the argument is, makes a very different semantic contribution compared to the BA pattern (or rather the B-part of the BA pattern). The RFR pattern conveys that there is an alternative to the proposition which remains open (see Constant (2012) for a related analysis which claims that no other alternative can be safely asserted). This meaning is of course closely related to Büring (1997)'s proposed meaning for contrastive topics, so in that sense the two proposals are similar, but

[3]Constant (2014)'s account is similar to Büring (1997) and would therefore make the same predictions.

according to Wagner (2008; 2012) this meaning contribution has in fact nothing to do with contrastive topics, and crucially the BA pattern lacks this meaning altogether.

The claim is that in utterance with a BA pattern, the first accent is realized differently just because another accented constituent is following. The rise, that according to Jackendoff (1972) crucially sets the B accent apart from the A accent, may simply be a continuation rise, or a related intonation cue for non-finality.

The prediction of Wagner (2008; 2012) is then that the responses both in (20) and (21) should be able to receive realizations with a BA pattern—provided of course that the non-focal constituent is accented at all. Under this view, the BA pattern does not in fact settle which constituent serves as the contrastive topic and which as the focus.

What then about Jackendoff (1972)'s examples with an AB pattern, with a final rise? According to Wagner (2012)'s account, this intonation pattern should only be used if a speakers intends to make it explicit that one or more alternatives remain open. In Jackendoff (1972)'s basic paradigm there is no reason why a speaker should do so, so we might not see many AB realizations. Also, to the extent that AB patterns are realized, the theory does not predict that this should be dependent on whether the first or second DP is a contrastive topic. In the following, we discuss two ways to implement this double focus account.

2.2.1 Simple Focus Account

The simplest way of implementing this double focus account is to treat the responses in both (20) and (21) as a case of double focus (see 24).

(24) Who ate what?

FRED$_F$ ate the BEANS$_F$

Note that this would not be possible in Jackendoff (1972)'s theory, where there is always only one primary focus. For Rooth (1992)'s theory of focus on the other hand this

is not problematic (see 25): One can have a focus operator (in this case the squiggle) that abstracts over both focused constituents.

(25) Who ate what?

$\sim$[FRED$_F$ ate the BEANS$_F$]

The antecedent for the presuppositional focus operator is provided by the wh-question. In these cases, the idea is that the first focus can be realized with a B-accent, and the second with an A-accent, simply because non-final accents can optionally be realized as B-accents, that is with a rise before the following accentual phrase. I will refer to this account as the Simple-Focus account. This approach can be made plausible by the observation that in response to multiple wh-questions, the order of the constituents can be freely swapped (26) while keeping the order of B and A accent constant. This is unexpected if these structures have radically different interpretations, as is predicted by Jackendoff (1972) and Büring (1997)'s analyses.

(26) Who bought what?

B: FRED bought the BEANS and the SPINACH was bought by SUE
B A B A

What about Jackendoff (1972)'s more complex contexts with a double wh-question plus a more local single wh-question? According to the simple-focus account, this more complex context provides two potential antecedents for focus marking: Picking the local antecedent will lead to deaccentuation of the non-focal constituent, picking the double-wh question will pick the double accented realization (27 and 28).

(27) Second constituent focused

[Who ate what?]$_i$ [What about Fred, what did he eat?]$_j$

a. Single focus: $\sim_j$ [Fred ate the beans$_F$]

b. Double focus: $\sim_i$ [Fred$_F$ ate the beans$_F$]

(28) First constituent focused

[Who ate what?]$_i$ [What about beans, who ate those?]$_j$

a. Single focus: $\sim_j$ [Fred$_F$ ate the beans]

b. Double focus: $\sim_i$ [Fred$_F$ ate the beans$_F$]

The Simple-Focus account predicts that to the extent that both constituents are focused and hence accented, the responses should sound the same in the two contexts, and in fact also the same to the simpler context with a simple multiple wh-question (see 29).

(29) Double Focus

[Who ate what?]$_i$

$\sim_i$ [Fred$_F$ ate the beans$_F$]

Under the Simple-Focus account, there is no such thing as prosodic contrastive topic marking, all there is, is a rise that differentiates a non final accentual phrase (B accent) from the final accentual phrase (A accent), which is followed by a fall. This is not to say that from an intuitive or functional way, one of the constituents does not serve the function of topic and the other that of a focus. It simply means that this distinction, to the extent that it is grammatically represented at all, does not affect prosody.

2.2.2 Nested Focus operators

The analysis in Wagner (2012) is almost that of the simple-focus account: Contrastive topics and foci are simply foci that associate with Rooth's $\sim$ operator. But there is an additional twist, based on a claim about the infelicity of certain word orders. The idea is that in sentence with contrastive topics, there are actually two $\sim$ operators involved, one with lower scope that associates with a single focus, and one with wider scope that associates with both foci (see 30).

(30) $\sim$ [[Fred]$_F$ $\sim$ [at the beans]$_F$]

The notion 'contrastive topic' plays no role in this account as such, but can be reconstructed as follows: In a configuration in which two $\sim$ operators are nested, foci that are only in the scope (and hence associates with) the higher operator are called contrastive topics, foci that are in the scope of both (and hence associate with both) are called foci.

The empirical claim that motivates focus-operator nesting is a purported linear order asymmetry. Wagner claims that contrastive topics (in the reconstructed sense here) cannot follow foci (see 31) therefore even though a B accent on the contrastive topic *Fred* would be fine in (31a), it is argued to be less acceptable with a topicalized syntax (31b).

(31) A: John ate the spinach.

A: What about Fred? What did he eat?

a. B: FRED ate the BEANS.
B A

b. # B: The BEANS, FRED ate.
A B

The analysis of this word order effect is as follows: In Wagner (2012)'s paradigm, both antecedents are actually used, by two different focus operators. An inner focus operator uses the immediate wh-question in the context, which operates then over alternatives just with the focused constituent; in addition there is an outer focus operator which operates over both foci, just like the double focus in the simpler account outlined above. Together with additional assumptions about scope this is intended to capture the observation that contrastive topics have to precede foci — contrary to Jackendoff (1972)'s and Büring (1997)'s claim.

2.3 Expectations

The experiments reported in this chapter were designed to test whether Jackendoff (1972)'s and Büring (1997)'s AB-BA hypothesis is correct, and furthermore whether the proposed word order preference reported in Wagner (2012) is attested. The predictions

of the various accounts are illustrated in table 2.1. It should be noted that the predictions shown in table 2.1 only account for cases in which there are two accented constituents. In cases where there is only one focus in the sentence (i.e. when the speaker interprets the sentence using the immediate wh-question), one expects more deaccentuations on the non-focal DP which would be the second DP in the *First DP focus* condition and the first DP for the *Second DP focus* condition. For the *Double focus* condition few deaccentuations are expected. This deaccentuation pattern is predicted by all accounts. With "deaccentuation" I mean that prominence is shifted (which may or may not involve complete deaccentuation of the non-prominent parts), moreover, what happens to the non-prominent part I call prosodic subordination.

Where it gets interesting are those cases in which both DPs are accented. Crucially for Jackendoff (1972) and Büring (1997) the order of the A and B accents is switched around depending on whether the first constituent is focused or the second, with the non-focal constituent considered to be the contrastive topic. In the double focus condition both AB and BA patterns are expected except for the AA pattern which is allowed by Büring (1997)'s account but not by Jackendoff (1972)'s account. Based on the Simple-Focus account, the same intonation pattern is expected across the different focus conditions: As long as there are two accented constituents the same pattern is expected regardless of the grammatical function of that constituent. Wagner (2012)'s Nested-Focus account makes similar predictions to the Simple-Focus account except that contrastive topics cannot follow foci and therefore the BA pattern is expected to be less acceptable in the *First DP focus* condition.

	Double focus	**First DP focus**	**Second DP focus**
Jackendoff (1972)	A-B/B-A	A-B	B-A
Büring (1997)	A-B/B-A/A-A	A-B	B-A
Simple-Focus	B-A/A-A	B-A/A-A	B-A/A-A
Nested-Focus	B-A/A-A	#B-A/A-A	B-A/A-A

Table 2.1: Expectations of double accented cases according to different focus accounts

To test the different focus accounts two experiments will be performed: Experiment 1.1 consists of two focus conditions (First DP focus and Second DP focus) and stimuli with explicit contrastive items in the context. The stimuli in this experiment are more or less analogous to the examples Jackendoff (1972) discusses and would therefore provide direct evidence in favor of or against Jackendoff (1972)'s and Büring (1997)'s claims. Experiment 1.2 has one extra focus condition, namely the double focus condition, furthermore the stimuli do not contain an explicit contrastive item in the context. The double focus condition serves as a control condition testing whether a sentence with double focus would be produced differently from sentences with 'contrastive topics' which could either be interpreted as a double focus construction or a single focus construction.[4]

2.4 Phonology accents

Before moving on to the actual experiment, let us first examine the phonological forms of the different accents that are important for the AB-BA patterns. This section provides a brief phonological description of the accents that are crucial to the BA and AB patterns and that have therefore been annotated in this experiment.

2.4.1 A accents

Two specific intonation patterns are of interest here, the so-called "A accent" and "B accent" which have been named so by Bolinger (1958).[5] Let us start with the A accent which is the least controversial one. Generally speaking the "A accent" is characterized by a high pitch on the accented syllable, followed by an abrupt fall in pitch towards the onset of the next vowel (see Figure 2.1). In fast speech the fall may be postponed to the second following syllable, but rarely beyond this (Bolinger 1958; Bolinger and Bolinger 1986).

[4]Focus here is exclusively determined by the local or global (implicit) wh-question in the context.
[5]It should be noted that Bolinger (1958)'s A and B accent do not correspond to Jackendoff (1972)'s A and B accents as we will see later on.

Following ToBI, the A accent will be represented as H*L, i.e. a high target followed by a low trailing tone.

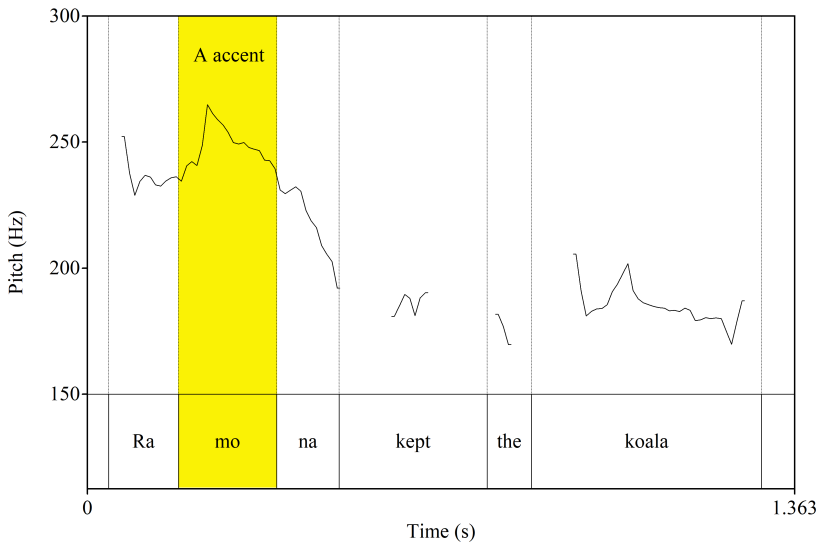


Fig. 2.1: Pitch trajectory “Ramona kept the koala" with an A accent on *Ramona*. The highlighted part in yellow marks the accented syllable. Click here to listen.

2.4.2 B accents

Jackendoff’s B accent: Rise-Fall-Rise

Generally speaking the “B accent" is realized as a fall in pitch, followed by a rise in pitch as can be clearly seen from Figure 2.2. The pitch peak is reached in the accented syllable and the rise ends at the end of the phrase. This B accent will be referred to as the “Rise-Fall-Rise (B) accent" and it is this description of the B accent that Jackendoff (1972) had in mind.

According to Ward and Hirschberg (1985) the Rise-Fall-Rise accent is known under a variety of names in various intonational frameworks: by Pike (1945) as a ’°2-4-3’ contour; by Bolinger (1958) as a subtype of ’Accent A’; by Kingdon (1958) as ’Tones III and V’; by Schubiger (1958), Gunter (1972) and Ward and Hirschberg (1985) as a ’falling-rising’ contour (or FR contour); by Halliday (1967) as ’Tone 4’; by Jackendoff (1972) as Bolinger’s ’B

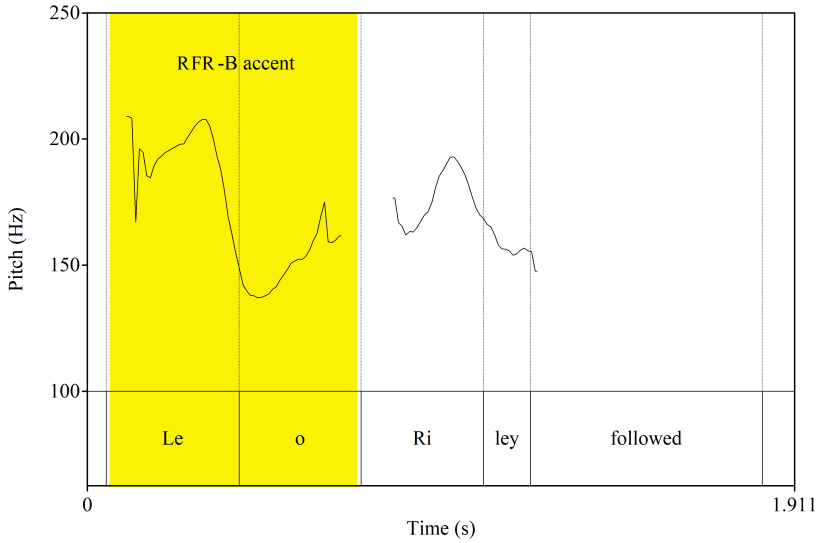


Fig. 2.2: Pitch trajectory 'Leo Riley followed" with a Rise-Fall-Rise (B) accent on *Leo*. The highlighted part in yellow marks the accented word. Click here to listen.

Accent'; by O'Connor and Arnold (1973), Ladd Jr (1980), and Cutler (1977) as 'fall rise'; by Liberman and Sag (1974) as 'contrastive stress within contradiction contour'; by Bing (1979) as an 'A-rise' contour; and by Moulton (1982) as '2°32↑'. This diversity of terminology reflects disagreement along two aspects: (a) what the phonetic properties of the contour are, and how it should be classified with respect to other intonational contours; and (b) how its 'meaning' should be characterized-and whether the contribution it makes to utterance interpretation is syntactic, semantic, or pragmatic (Ward and Hirschberg 1985).

Important to notice is that what Jackendoff (1972) calls Bolinger's B accent is in fact a subtype of Bolinger's 'Accent A' according to Ward and Hirschberg (1985). However Bolinger's subtype of the 'Accent A' is fundamentally different from Jackendoff (1972)'s B accent in that it lacks a terminal rise as can becomes clear from Bolinger (1958)'s description of this sub-type A accent: "One rather sharp deviation, which for semantic reasons I would class as a sub-type of A, puts the accentable syllable at a lower pitch than the one immediately following, but requires that only that one weak syllable remain high, the syllable after it must come down rapidly." (p. 143). The sub-type A accent can be seen as an A accent with a delayed peak, i.e. the peak either occurs late in the accented syllable or

in the following syllable, thus resulting in an apparent Rise-Fall gesture. In other words it is an A accent which is aligned later in the accented syllable. Bolinger (1958)'s B accent also does not completely correspond to Jackendoff (1972)'s B accent since it only describes a rising accent, thus lacking the falling part altogether. Therefore it can be concluded that Bolinger (1958)'s B accent has little to do with Jackendoff (1972)'s B accent, even though Jackendoff (1972) ascribes this B accent to Bolinger.

In terms of phonological representation Ward and Hirschberg (1985) represents the Jackendoff (1972)'s B accent as L*HLH% (or L*+H-L-H%), i.e. an L target followed by an H and L trailing tones and an H final boundary tone. Ward and Hirschberg (1985) explicitly argues against a representation with an H target tone like LH*LH% (or L- +H*L-H%) claiming that it is phonologically a different contour. In this study however, we will represent the RFR contour as LH*LH%, i.e. a low leading tone followed by a H target on the accented syllable followed by an L trailing tone which spreads until the H final boundary tone. The reason for choosing an H target rather than an L target as proposed by Ward and Hirschberg (1985) is that the initial L tone seems to be phonetically reduced especially in phrase-initial position (see 2.2).

Bolinger's B accent: Rise and Early Rise (B) accent

In this section we will briefly describe two more pitch accents which can be considered a B accent, namely the 'Early-Rise' and 'Rise' accent. The 'Early-Rise' pitch accent is characterized by a high pitch on the accented syllable, but unlike the A accent a falling gesture is not part of this pitch accent. In ToBI this would be either transcribed as H* or LH*.

The distinction between H* and LH* is one of the most contentious in the ToBI framework (Calhoun 2012). Several major studies investigating inter-transcriber reliability have concluded that the L+H*/H* distinction was so difficult to make that these two categories had to be collapsed and agreement could not be reported (Pitrelli et al. 1994; Silverman et al. 1992; Syrdal and McGory 2000). The issue is that with sufficient preceding unac-

cented syllables, the F0 tends to fall before an accentual rise, this apparent low can be hard to distinguish from an L target (Ladd 2008; Ladd et al. 2009). I have chosen to represent the early rise accent as H* this is based on the observation that in the pre-accentual part it may either come from an equally high level as is the case in Figure 2.3 or from a lower level which would result in an apparent pre-accentual rise (see Figure 2.4): Crucially the accented syllable itself is perceived as high.

The rise accent, on the other hand, is characterized by a rising gesture on the accented syllable (see Figure 2.5). The accented syllable starts low and the peak occurs at the end of the accented syllable or in the following syllable, it is transcribed as L*H in ToBI. It should be noted that the rise and early rise accents are often considered B accents or at least the intuition is that they sound similar to Rise-Fall-Rise B accents (perhaps a more phonetically reduced form of a B accent) (Calhoun 2012). In fact, as we have seen in the previous section, Bolinger (1958)'s original B accent actually describes a rising accent and other researchers may actual have Bolinger (1958)'s rising accent in mind when discussing the B accent rather than Jackendoff (1972)'s Rise-Fall-Rise B accent. Therefore, it will be essential to keep in mind that there are two B accents: Jackendoff (1972)'s B accent, i.e. a rise-fall-rise B accent and Bolinger (1958)'s B accent, i.e. a rising B accent. According to Wagner (personal communication, March 4, 2022), it is Bolinger (1958)'s description of a B accent which is usually meant when referring to a B accent or at least the intuition is that they sound similar to Rise-Fall-Rise B accents. In this study when using the term "B accent" (without specifying which specific B accent is referred to) any B accent is meant (i.e. Rise-Fall-Rise, Early Rise and Rise accent).

2.4.3 Metrical Structure

Recall from the previous chapter that the metrical structure of a sentence can among other factors be determined by focus. A wh-question determines which constituent receives focus, which results in that constituent becoming prominent. Based on this assumption, in a

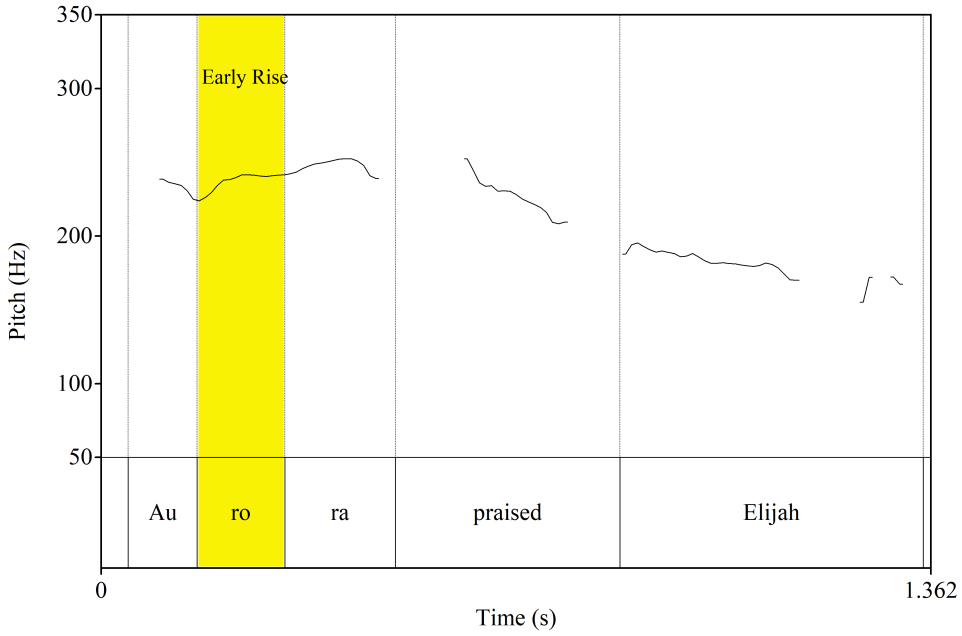


Fig. 2.3: Pitch trajectory 'Aurora praised Elijah" with a Early-Rise accent on *Aurora*. The highlighted part in yellow marks the accented syllable. The Early Rise accent comes from a high onset and therefore no rise is observed. Click here to listen.

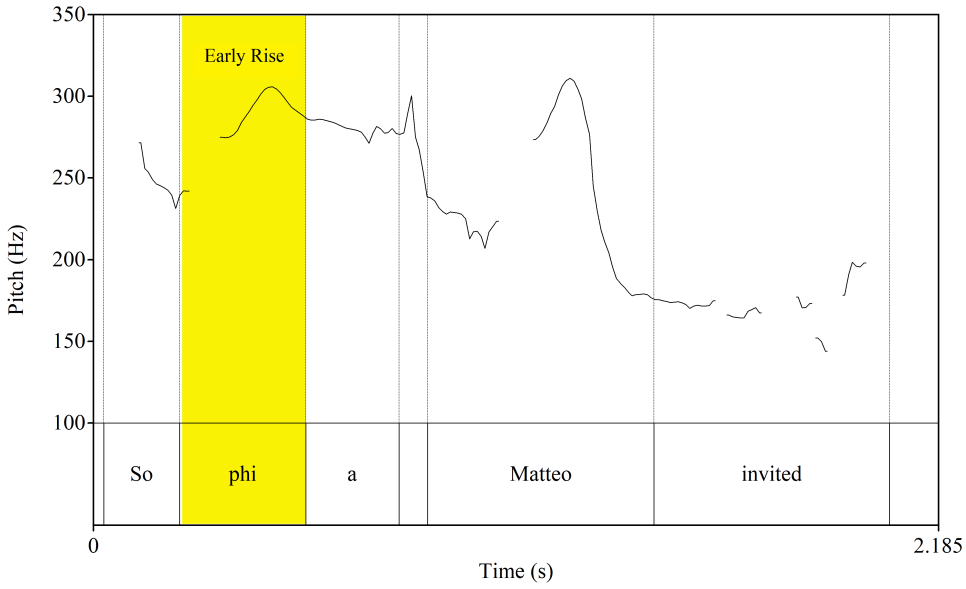


Fig. 2.4: Pitch trajectory 'Sophia Matteo invited" with an Early Rise accent on *Sophia*. The peak of the Early Rise accent occurs within the accented syllable. There is a pitch reset after the Early-Rise pitch accent and an A accent on *Matteo*.Click here to listen.

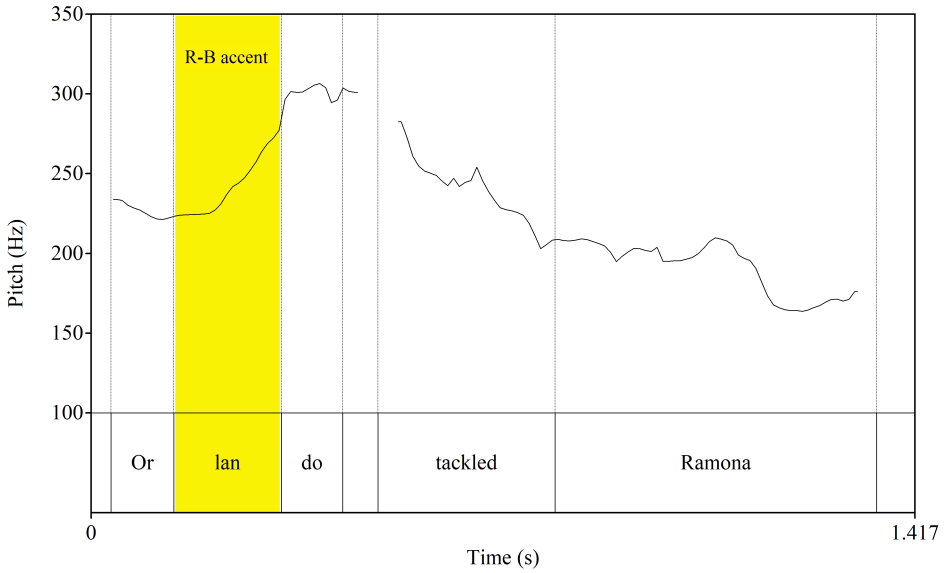


Fig. 2.5: Pitch trajectory 'Orlando tackled Ramona" with a Rise accent on *Orlando*. Click here to listen.

dialogue like Example (20), *beans* would receive focus (and therefore be considered prominent) since the wh-word is directly questioning what was eaten. In active sentences this means that focus occurs sentence-finally (see Example 32a), but in passive or topicalized sentences a prominence shift is expected (see Examples 32b and 32c). The inverse metrical pattern is expected for a sentence with subject focus, see Example (33). Constituents that are not focused would be considered deaccented and would therefore not receive a pitch accent.

Based on the assumption that focused constituents are accented and thus realized with a pitch accent, the expectation is that in the object/subject focus conditions there is only one pitch accent, namely on the focused element, all other elements are deaccented and receive no pitch accent. In the double focus condition both DPs are focused and would therefore receive a pitch accent (see 34). Although the metrical structure of a sentence plays a crucial role in determining where pitch accents can occur there are other mechanisms that can determine whether a constituent is accented or not and therefore whether it would receive a pitch accent. It should be noted that Jackendoff (1972) only discusses those cases in which both the topic and focus are accented.

How then does one end up with a double accent realization in the subject and object focus conditions? According to Jackendoff (1972) it is those contexts which make a certain constituent to be interpreted as a contrastive topic that can optionally be produced with both constituents accented. The context itself does not provide an explicit multiple wh-question (like in the double focus condition), but makes it implicitly clear that such a broader question is relevant as we have seen in previous sections.

(32) A: Well, what about Fred? What did he eat? (Object Focus)

a. B: FRED ate the BEANS. (active syntax)
x (focus)

b. B: The BEANS were eaten by FRED. (passive syntax)
x (focus)

c. B: The BEANS Fred ate. (topic syntax)
x (focus)

(33) A: Well, what about the beans? Who ate those? (Subject Focus)

a. B: FRED ate the beans. (active syntax)
x (focus)

b. B: The beans were eaten by FRED. (passive syntax)
x (focus)

c. B: The beans FRED ate. (topic syntax)
x (focus)

(34) A: Who ate what? (Double Focus)

a. B: FRED ate the BEANS. (active syntax)
x x (focus)

b. B: The BEANS were eaten by FRED. (passive syntax)
x x (focus)

c. B: The BEANS FRED ate. (topic syntax)
x x (focus)

It is important to note that the metrical structure itself does not functionally distinguish contrastive topics from foci, it merely tells us how many prominent syllables there are in a certain sentence and how prominent they are relative to each other. Given a double

accented realization Jackendoff (1972) and Büring (1997) expect B accents to align with the contrastive topic and A accents with the focus of the sentence (see Example 35). According to the simple-focus account there is no link between the function of a constituent in a sentence and the pitch accent it would align with, instead B accents can be optionally used to indicate that it is a non-final accent. Contrary to the simple-focus account, Wagner (2012) predicts BA patterns to be less felicitous in contexts like (35) since a nested focus structure would result in a interpretation of the sentence that does not match its actual meaning, rather a non-nested double focus structure would be expected in such cases which would result in an AA pattern.

(35) A: Well, what about the beans? Who ate those? (Subject Focus)

B: FRED ate the BEANS.

x	x	(focus)
A	B	(Jackendoff 1972; Büring 1997)
B	A	(Simple-Focus)
A	A	(Wagner 2008; 2012)

2.5 Experiment 1.1

In order to test the different focus accounts and more specifically the claims made by Jackendoff (1972), a production experiment was set up using MatLab. Participants were presented, both in written form as well as in auditory form, with a number of carefully manipulated questions, after which they were asked to record their answer and assess how well their answer sounded. Crucially, the stimuli in this experiment contain a contextually given explicit contrast to the DPs in the answer.

2.5.1 Participants

A total of 25 participants finished the experiment, all of them were native speakers of North-American English. A compensation of 3.00 C$ was given which boils down to an hourly rate of 12.00 C$ since the experiment itself took about 15 minutes.

2.5.2 Stimuli

Each stimulus consists of a one-sentence context, two questions and an answer. There are two different focus conditions, namely the object focus and subject focus. The *object focus* condition consists of a wh-question questioning the object (see 36a) and the *subject focus* condition consists of a wh-question questioning the subject (see 36b). Each question evoking focus is preceded by a "*What about...?* question which is argued to evoke a contrastive topic. The context sentence preceding both questions consists of two explicit contrastive constituents, in this case *Amy* and *Danny*, where *Amy* contrasts with *Nolan* and *Danny* with *Morgan* (see Example 36). The questions have been prerecorded by the researchers and serve as an auditory stimuli to be presented to the participants during the experiment. A more detailed description of how these elements are presented to the participant is given in the *procedure* section. An example of such a stimuli set is provided below[6]:

(36) Questions:

a. I know that Amy criticized Danny. But what about Nolan? Who did he criticize? (Object focus)

b. I know that Amy criticized Danny. But what about Morgan? Who criticized him? (Subject focus)

For the answers the word order has been manipulated, there are three different word orders: Active, passive and topic word order (see 37). What is crucial for these sentences is the order of the object and subject of the sentence. Depending on the question either the subject or object is focused while the other constituent would constitute the contrastive topic.

(37) Answers:

a. Nolan criticized Morgan. (Active Syntax)

b. Morgan was criticized by Nolan. (Passive Syntax)

[6] All stimuli sets can be found in the *spreadsheets* folder on the OSF-page. Click here.

c. Morgan Nolan criticized. (Topic Syntax)

For analysis purposes the focus conditions have been relabeled to *First DP focus* and *Second DP focus*. The first DP is focused in the object focus condition for sentences with a passive and topicalized syntax, and in the subject focus condition for sentences with an active syntax (*First DP Focus*). The inverse is true for *Second DP focus*: The second DP is focused in the subject focus condition for sentences with a passive and topicalized syntax, and in the object focus condition for sentences with an active syntax (*Second DP focus*). In total there are six conditions, i.e. two question conditions and three answer conditions. 36 stimuli sets were created such that each condition would be repeated six times.

Based on Jackendoff (1972) and Büring (1997) an interaction effect is expected between the different focus conditions (*focus*) and whether the accent occurs on the first DP or second DP (*linear order*): For A accents most are expected to be produced on the second DP in the *Second DP focus* condition and on the first DP in the *First DP focus* condition. For B accents (i.e. Rise-Fall-Rise (B) accent according to Jackendoff (1972)) the opposite pattern is expected.

The double focus accounts predict no interaction effect between *focus* and *linear order*, instead a main effect of *linear order* is expected according to which more B accents (i.e. Rise-Fall-Rise (B), Early Rise (B) or Rise (B) accent) would be produced on the first DP compared to the second DP and a similar but opposite trend is expected for A accents (i.e. more A accents on the second DP than on the first DP). Different from the simple-focus account, Wagner (2012) predicts a main effect of *focus* according to which no BA patterns should be produced in the *First DP focus* condition and if they are produced these patterns should be less acceptable.

2.5.3 Procedure

Each participant had to perform 36 trials which were presented to them in a latin-square design. Participants were seated in front of a screen in a sound-proof room in the Depart-

ment of Linguistics at McGill University. Each trial starts off with a fixation point, this is to ensure the participant knows exactly when one trial ends and another one starts. Next, the context and question are provided in written form: the context ensures the phrases sound more natural rather than out of the blue (see the stimuli section for an actual example).

On the next screen, the context and question are presented auditorily. As soon as the question ends the participant is instructed to record the answer as natural as possible. The trial finishes off with an acceptability input screen in which the recorded answer is played back to the participant and he/she has to indicate how natural their answer sounded based on a scale from 1 to 8 where 1 means completely unnatural and 8 means completely natural.

2.5.4 Statistical Analyses

A total of 900 recordings were made (25 participants*36 trials = 900), one participant had to be excluded since no recordings were made and an additional 14 recordings had to be excluded due to poor sound quality. Eventually, a total of 850 recordings were obtained for data analysis. The recorded sentences were annotated by the author for the pitch accents on the first DP and second DP. The following categories were distinguished for the annotation: Falling (A) accent, Rise-Fall-Rise (B) accent, Early Rise (B) accent, Rise (B) accent and Deaccented. Whenever there was perceived to be a lack of pitch accent it was annotated as *Deaccented*. For a phonological description of each pitch accent see the respective section earlier in this chapter.

Figure 2.6 shows the proportions of pitch accents in percentage for each condition. There is noticeably more deaccentuation on the second DP when the first DP is focused, but the inverse is not observed, i.e. more deaccentuation on the first DP when the second DP is focused. This asymmetry is unexpected and I will come back to this asymmetry at the end of this chapter. Jackendoff (1972) is mainly interested in cases where both DPs are accented, i.e. without deaccentuation, therefore Figure 2.7 where all cases with deaccen-

tuation have been excluded, is more informative: the exact numbers can be found in table 2.2 in Appendix A. Of the 850 recordings, 573 recordings contained no deaccentuations and were used for the statistical analyses.

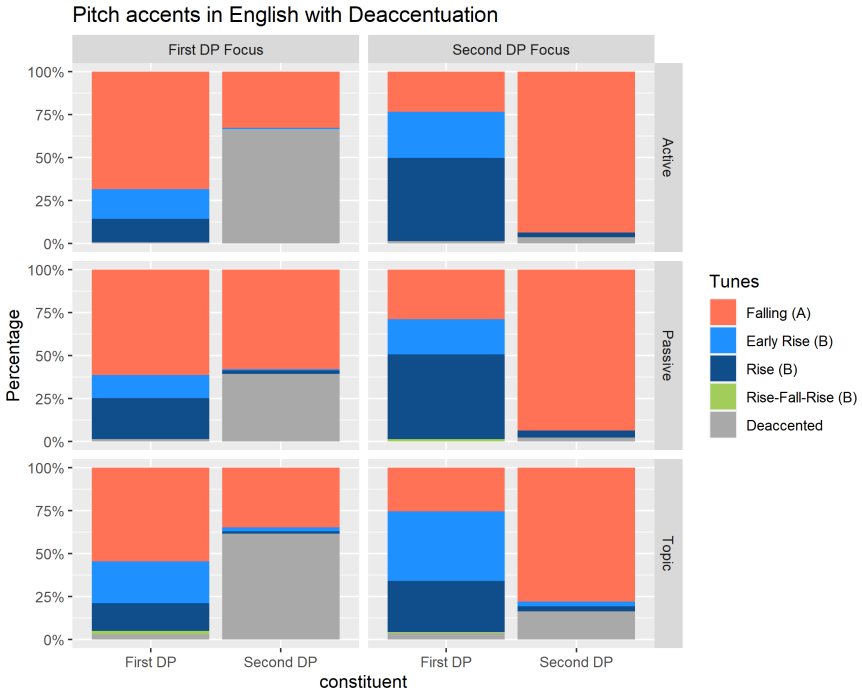


Fig. 2.6: Experiment1.1: Pitch Accents with Deaccentuation in English

To make more concrete conclusions several statistical analyses were performed on the proportion data corresponding to the different pitch accents annotated per condition. The dependent variable is the presence or absence of a certain pitch accent, where the absence of a pitch accent corresponds to the pooled group consisting of all other pitch accents not present: E.g. in the proportion data of the Falling (A) accent, present means it is a Falling (A) accent and absence corresponds to any pitch accent but the A accent which in this case would correspond to any B accents (i.e. Rise-Fall-Rise, Early Rise and Rise B accents).

Although the data is binary (i.e. absence/presence of a pitch accent) and a binomial logistic regression is often considered most appropriate for binary data, I have chosen to perform a (modified) poisson regression instead based on conclusions drawn by other studies. The estimates obtained by a poisson regression are more conservative according

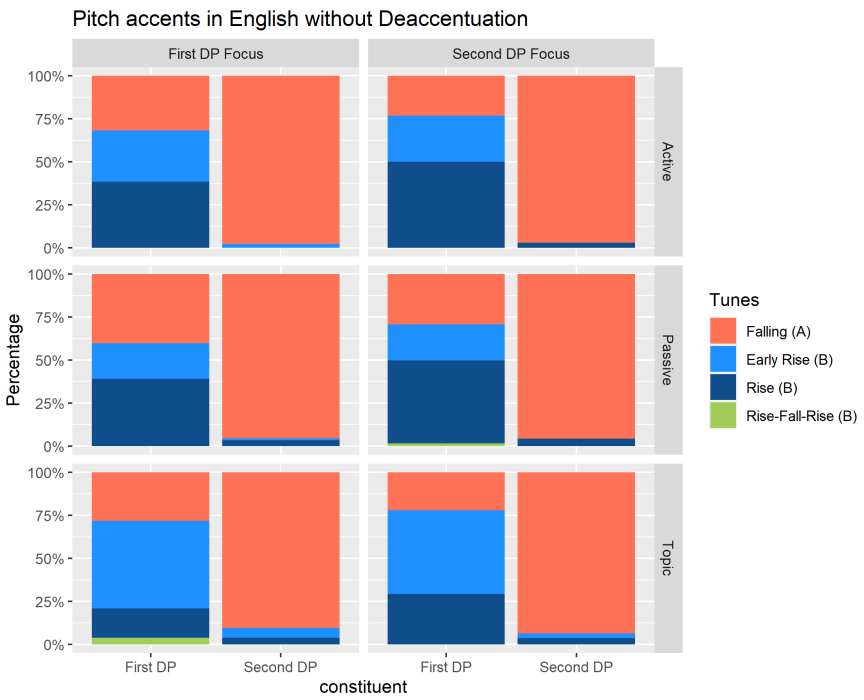


Fig. 2.7: Experiment1.1: Pitch Accents without Deaccentuation in English

to Cleophas and Zwinderman (2016), in addition estimates are relatively robust to omitted variables in contrast to logistic regression (Zou 2004). A regular poisson regression assumes equal mean and variance of the data, for the proportion data in this study however the mean is often larger than the variance (i.e. there is underdispersion). To account for this underdispersion I follow Zou (2004) by using a modified Poisson regression approach (i.e., Poisson regression with a robust error variance[7]).

2.5.5 Results

Figure 2.8a gives a visualization of the distribution of the *A accent* across different conditions. From this figure it becomes clear that more A accents were produced on the second DP than on the first DP regardless of the word order. To see if there are any significant results a modified poisson regression was performed using R (Team 2008) and the glm function in the MASS Rpackage (Jackman et al. 2015). A poisson regression model was

[7]The robust error variance procedure used is known as sandwich estimation which can be found in the *sandwich* Rpackage

fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *A accents* was taken as independent variable and as dependent variables *focus*, *syntax* and *linear order* were included in the model.[8] The final model performed significantly better than an intercept-only base line model (χ^2(4): 221.76, p = 0) and is a good, optimal fit (C: 0.8733344, Somers' Dxy: 0.7466689). The model's explanatory power is substantial (Nagelkerke's R2 = 0.39).[9] From the final model it becomes clear that only the *linear order* has a significantly effect: More A accents were produced on the second DP than on the first DP (beta = 1.22, 95% CI [1.04, 1.39], p < .001). No other significant effects nor interaction effects were reported.

The distribution of the Rise-Fall-Rise (B) accents across the different conditions is shown in Figure 2.8b. Two very general observation can be made, first of all very few rise-fall-rise accents have been produced and secondly those accents only occur on the first DP. A poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *Rise-Fall-Rise (B) accents* was taken as independent variable and as dependent variables *focus*, *syntax* and *linear order* were included in the model. The final model did not perform significantly better than an intercept-only base line model (χ^2(4): 9.045, p = 0.05998). No significant results were found for this pitch accent, this may be due to the low number of Rise-Fall-Rise (B) accents produced.

The distribution of the rising B accent across the different conditions is shown in Figure 2.9a. From this figure it becomes clear that in general more rise accents were produced on the first DP than on the second, and it seems that the least number of rise accents was produced in the topic word order. A mixed-effects poisson regression model[10] with random intercepts for speakers was fitted to the data in a step-wise-step up procedure. In

[8]See Appendix B for the statistical models.

[9]The model's intercept, corresponding to focus = First DP Focus, constituent = First DP and syntax = Active, is at -1.22 (95% CI [-1.45, -1.01], p < .001).

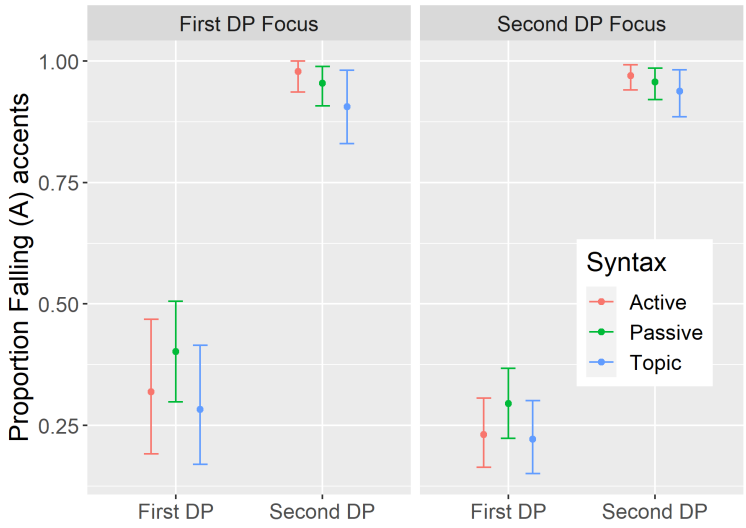
[10]Sometimes a mixed-effects poisson regression was performed and sometimes a regular poisson regression was performed. When a random effects structure did not significantly improve the model, a regular poisson regression was run.

the final minimal adequate model the proportion of *R-B accents* was taken as independent variable and as dependent variables *focus*, *syntax* and *linear order* were included in the model. The final model performed significantly better than an intercept-only base line model (χ^2(4): 220.39, p = 0) and is a good, optimal fit (C: 0.906126, Somers' Dxy: 0.812252). The model's total explanatory power is substantial (conditional R2 = 0.53) and the part related to the fixed effects alone (marginal R2) is of 0.39.[11] From the final model it becomes clear that the effect of *linear order* is significant: Less R-B accents were produced on the second DP than on the first DP (beta = -2.48, 95% CI [-2.95, -2.02], p < .001). *Focus* also had a significant effect: More R-B accents were produced in the *Second DP focus* condition than in the *First DP focus* condition (beta = 0.34, 95% CI [0.06, 0.63], p = 0.019).

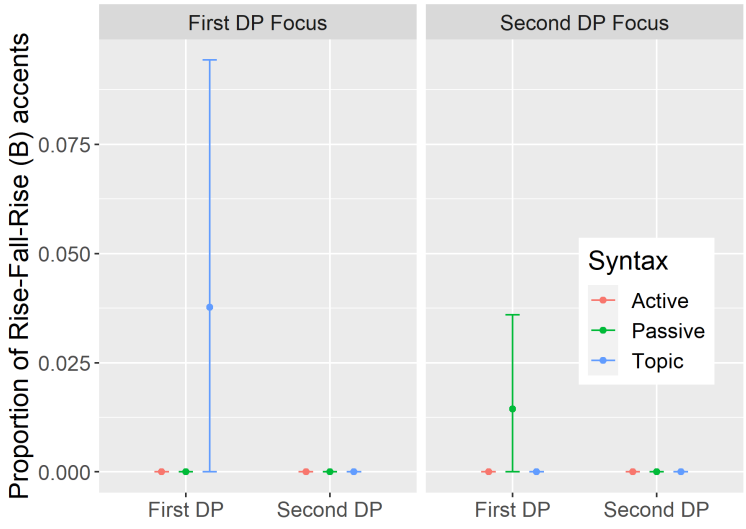
Finally, the distribution of the *Early Rises* across the different conditions is shown in Figure 2.9b. Once again there are more early rises on the first DP than on the second, but this time most early rises seem to have been produced in the topic word order condition. A mixed-effects poisson regression model with random intercepts for speakers was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *ER-B accents* was taken as independent variable and as dependent variables *syntax* and *linear order* were included in the model. The final model performed significantly better than an intercept-only base line model (χ^2(9): 217.25, p = 0) and is a good, optimal fit (C: 0.8900202, Somers' Dxy: 0.7800405). The model's total explanatory power is substantial (conditional R2 = 0.57) and the part related to the fixed effects alone (marginal R2) is of 0.52.[12] A significant effect of *linear order* was found: Less Early Rises were produced on the second DP than on the first DP (beta = -3.11, 95% CI [-3.81, -2.41], p < .001). In addition, there was a significant effect of *syntax*: Overall more early rises were produced in the *topic* word order than in the other word orders (beta = 0.55, 95% CI [0.20, 0.89], p = 0.002)

[11] The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = First DP Focus, is at -1.37 (95% CI [-1.82, -0.92], p < .001).

[12] The model's intercept, corresponding to constituent = First DP and syntax = Active, is at -1.40 (95% CI [-1.74, -1.06], p < .001).

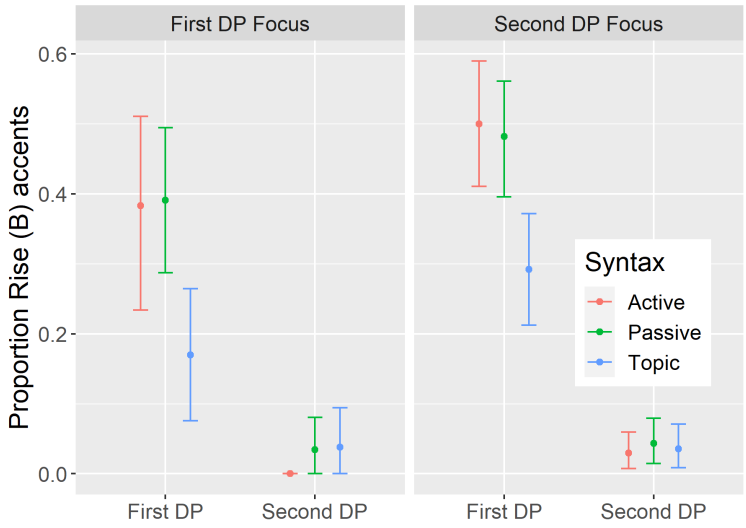


(a) A accents

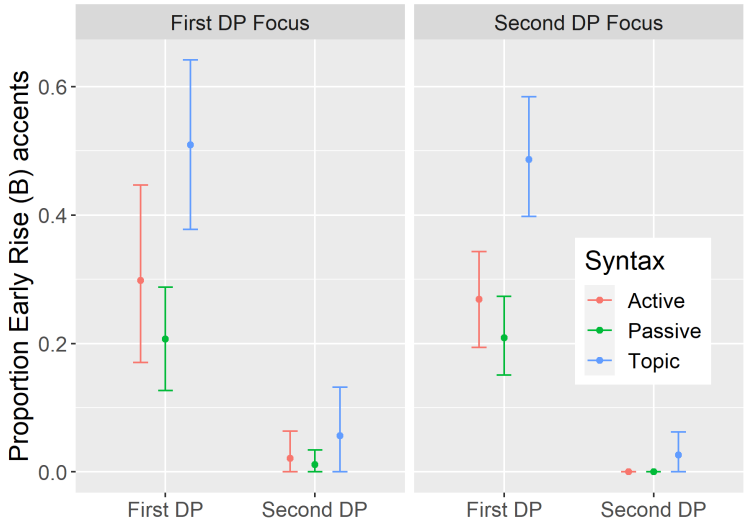


(b) Rise-Fall-Rise (B) accents

Fig. 2.8: Experiment 1.1: Distribution A and Rise-Fall-Rise (B) accents



(a) Rise (B) accents



(b) Early Rise (B) accents

Fig. 2.9: Experiment 1.1: Distribution Rise and Early Rise B accents

2.5.6 Conclusion

The results from Experiment 1.1 do not support Jackendoff (1972)'s or Büring (1997)'s approach to AB-BA patterns in English: The AB-BA pattern does not convey whether the B-accented or the A-accented functions as a contrastive topic in the utterance. There are very few actual Rise-Fall-Rise (B) accents and the few Rise-Fall-Rise accents that have been

produced occur consistently on the first DP, something completely unexpected based on predictions made by Jackendoff (1972). However as already mentioned in the section on the phonology of various pitch accents, this strict definition of a B accent may not actually correspond to what most researchers considered a B accent: Both the early rise and rising accent could be considered a B accent by those researchers who follow Bolinger (1958) original definition of a B accent. Even with this looser definition of a B accent there is no evidence in support of Jackendoff (1972)'s analysis, according to this analysis there should have been more rising pitch accents on the second DP in the *First DP Focus* condition and one would have expected an interaction effect between *focus* and *linear order*: None of the pitch accents showed an interaction effect.

The results do seem to support a *double focus analysis* and more specifically the simple-focus account according to which the type of pitch accent is determined by the linear order in which the pitch accents occur rather than the actual information status of the constituent with which it is aligned: A rising accent indicates that it is not the last accent and another one will follow. Additionally, Wagner (2012)'s nested account is only partially supported by the results. According to this account there should be no BA patterns in the *First DP Focus* since the nested focus analysis does not allow for contrastive topics to follow foci. Although it seems that there are slightly more rise (B) accents in the *Second DP Focus* (albeit not significant) as can be seen from Figure 2.7, there are still a substantial number of rising B accents produced in the *First DP Focus* condition.

More evidence against Wagner (2012)'s nested foci account comes from the acceptability judgements given to the participants' own productions. Despite the fact that BA patterns were produced in the *First DP Focus* condition, one could still argue in favor of Wagner (2012)'s account if these productions were actually assessed less acceptable compared to the other focus condition. Figure 2.10 however, tells us that this is not the case: In general the topic sentences are judged least acceptable, followed by the passive ones, but crucially no difference across the different focus conditions can be observed at least

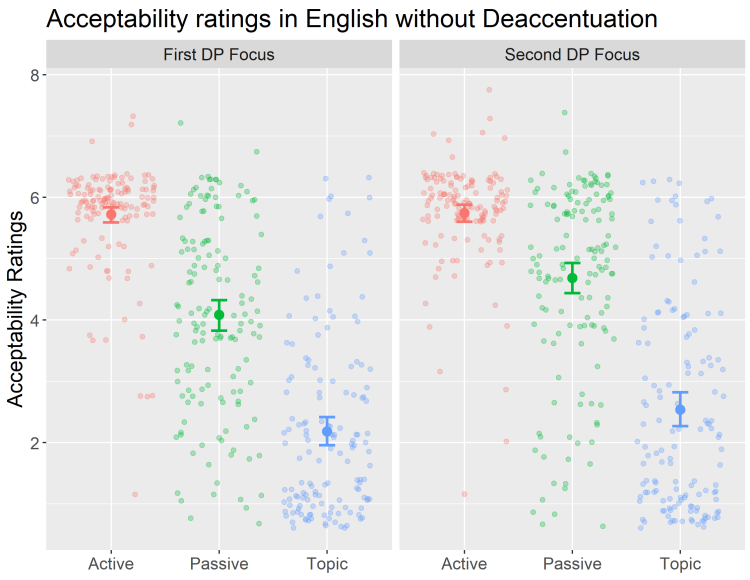


Fig. 2.10: Experiment 1.1: Acceptability Judgements

for the active and topic cases. The passive sentences do seem to display the pattern expected by Wagner (2012), i.e. lower acceptability ratings for the *First DP Focus* condition compared to the other focus condition in sentences with both DPs being accented. It is not completely clear to me why this would be the case, but this pattern seems to have more to do with the passive syntax behaving differently from other word orders rather than an overall lower rating for the *First DP Focus* condition.

One shortcoming of Experiment 1.1 is that it did not include simple cases where a context invites a response with double focus. Having an extra focus condition that explicitly evokes a double focus, by creating for instance a context containing a multiple-wh question, could in a sense serve as a control condition and provide more conclusive evidence in favour of a double focus analysis. Following the simple-focus analysis the expectation would be that the intonation pattern for the double focus condition should not be any different from the other conditions when realized with a double accent. Experiment 1.2 does precisely this, the setup is essentially the same as Experiment 1.1 but an extra *Double Focus* condition was added as a manipulation.

2.6 Experiment 1.2

Similar to Experiment 1.1, Experiment 1.2 aims to test the claims made by Jackendoff (1972). To do so an online production experiment was set up using jsPsych (de Leeuw 2015). The methodology used is roughly the same as in Experiment 1.1, therefore only those parts where the two experiments differ have been written down. One crucial difference is that an extra focus condition has been added, namely the double focus condition, this condition serves more or less as a control condition. The procedure is similar to Experiment 1.1 with the exception that participants are directed to the experiment via an online link rather than having to come to the lab.

2.6.1 Participants

A total of 30 participants were recruited through Prolific, all of them were native speakers of North-American English. A compensation of 4.50 C$ was given which boils down to an hourly rate of 13.50 C$ since the experiment itself took about 20 minutes.

2.6.2 Stimuli

Each stimulus consists of a question and an answer, there are three different focus conditions (see Example 38): double focus, object focus and subject focus. The *double focus* condition consists of a multiple wh-question (see 38a), the *object focus* condition consists of a wh-question questioning the object (see 38b) and the *subject focus* condition consists of a wh-question questioning the subject (see 38c). Each question evoking focus is preceded by a "*What about...?* question which evokes a topic. Different from Experiment 1.1 is the extra focus condition and the fact that there is no explicit contrastive item in the context. The reason for not having any explicit contrastive items is to keep the contexts as similar as possible across the different focus conditions. Because a contrastive context

does not work for the double focus condition, the choice has been made to use this general one-sentence context for all focus conditions.

(38) I heard the students were very critical of each other.

a. Who criticized whom? (Double focus)

b. What about Amy? Who did she criticize? (Object focus)

c. What about Nolan? Who criticized him? (Subject focus)

Similar to Experiment 1.1, for the answers there are three different word order: Active, passive and topic word order. What is crucial for these sentences is the order of the object and subject of the sentence. Again, for analysis purposes the focus conditions have been relabeled to *Double focus*, *First DP focus* and *Second DP focus*.

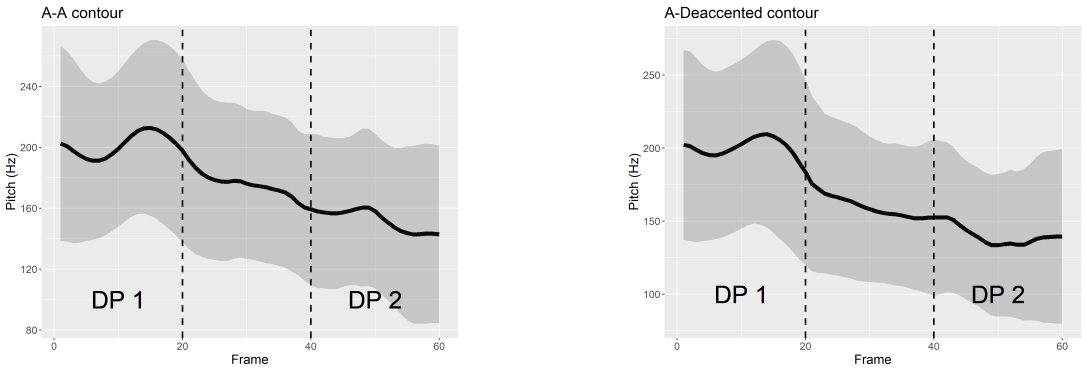
The predictions are similar to the ones for Experiment 1.1: Jackendoff (1972) and Büring (1997) expect an interaction effect between the different focus conditions and the *linear order* in which they occur: For A accents most are expected to be produced on the second DP in the *Second DP focus* condition and on the first DP in the *First DP focus* condition. For B accents the opposite pattern is expected. For the *Double focus* condition Jackendoff (1972) expects both BA and AB patterns to be produced but no AA pattern since each sentence can only contain one primary focus. Büring (1997) on the other hand expects any of the three patterns to be produced in the double focus condition depending on whether the sentence is interpreted as having a contrastive topic on either one of the DPs (AB or BA) or whether it is simply interpreted as a double focus construction (AA).

The simple-focus account predicts no interaction effect between *focus* and *linear order*, instead a main effect for *linear order* is expected according to which more B accents would be produced on the first DP compared to the second DP and vice versa for A accents (i.e. more A accents on the second DP than on the first DP). For Wagner (2012) an additional main effect of *focus* is predicted according to which no BA patterns should be produced in the *First DP focus* condition and if they are produced these patterns should be less

acceptable. Both double-focus accounts also expect AA patterns to be produced across the different focus conditions.

2.6.3 Results: Mean Pitch Trajectory

To give an idea of the actual productions, the mean pitch trajectories of some of the most frequent contours have been illustrated in this section (see Figures 2.11 and 2.12). For the exact numbers see Table 2.4 in the Appendix. Besides the most common contours, the RFR-A and A-RFR patterns are provided as well since these were the ones most expected by Jackendoff (1972) (see Figure 2.13), however there are only few instances of these contours.[13]

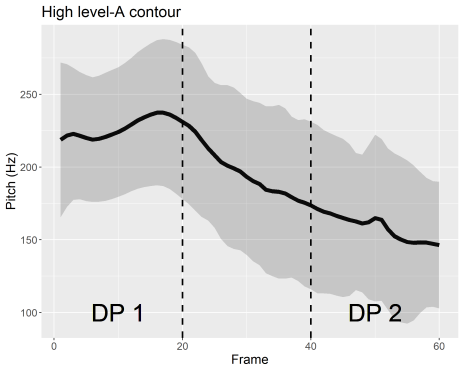


(a) A-A pattern (H*L H*L L%)

(b) A-Deaccented pattern (H*L L%)

Fig. 2.11: Mean pitch trajectories of active and passive sentences. The black line indicates the mean pitch and the shaded area the standard deviation.

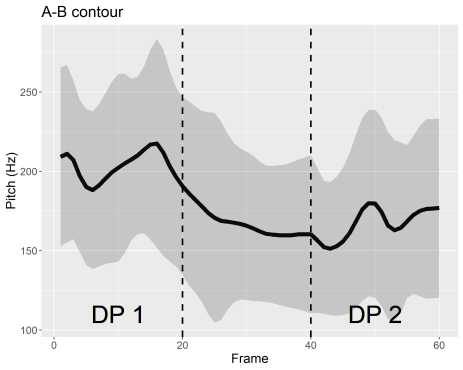
[13] More visualizations of mean pitch trajectories can be found can be found in the *data_analysis/plots/contours* folder on the OSF-page. Click here.



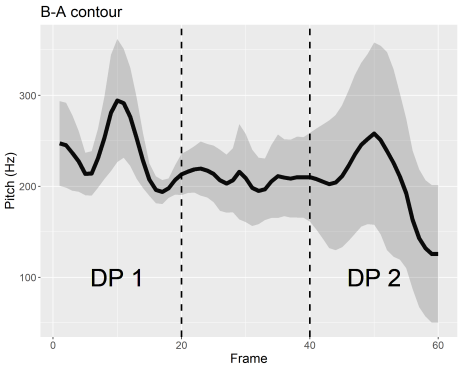
(a) Early Rise-A pattern (H* H*L L%)

(b) R-A pattern (L*H H*L L%)

Fig. 2.12: Mean pitch trajectories of active and passive sentences.



(a) A-RFR pattern (H*L LH*L H%)



(b) RFR-A pattern (LH*L H% H*L L%)

Fig. 2.13: Mean pitch trajectories of active and passive sentences.

2.6.4 Results

A total of 1080 recordings were made (30 participants*36 trials = 1080), one participant had to be excluded since this person was non-native and an additional 26 recordings had to be excluded due to poor sound quality, leaving us with 1018 good recordings. Eventually, 627 out of 1018 recordings contained no deaccentuations and were used for data analysis.

Figure 2.14 shows the proportions of pitch accents in percentage for each condition. As in the previous experiment there is noticeably more deaccentuation on the second DP when the first DP is focused, but the inverse is not observed, i.e. more deaccentuation on the first DP when the second DP is focused. For our purposes Figure 2.15 where all cases with deaccentuation have been excluded, is more informative.

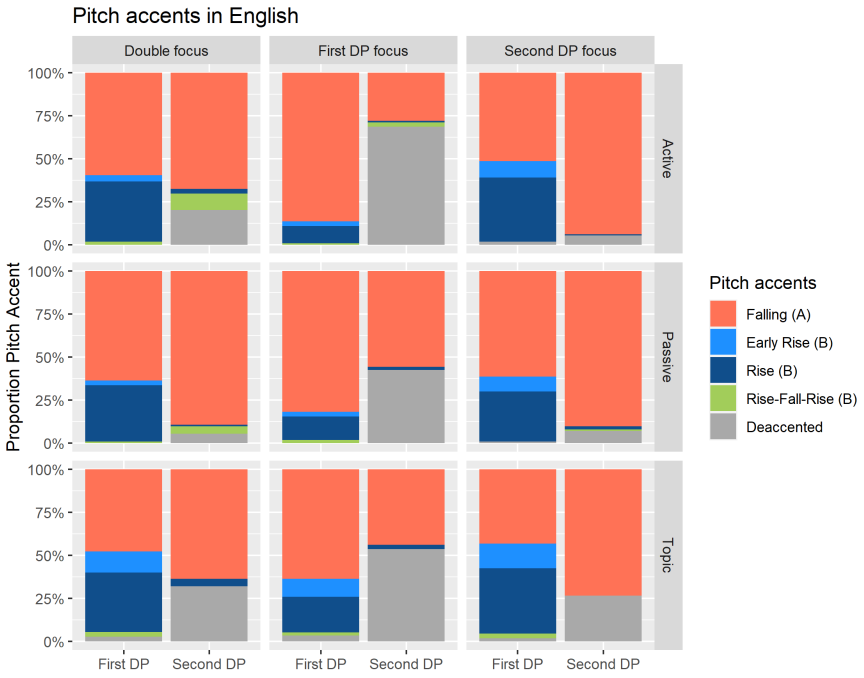


Fig. 2.14: Experiment 1.2: Pitch Accents with deaccentuation in English

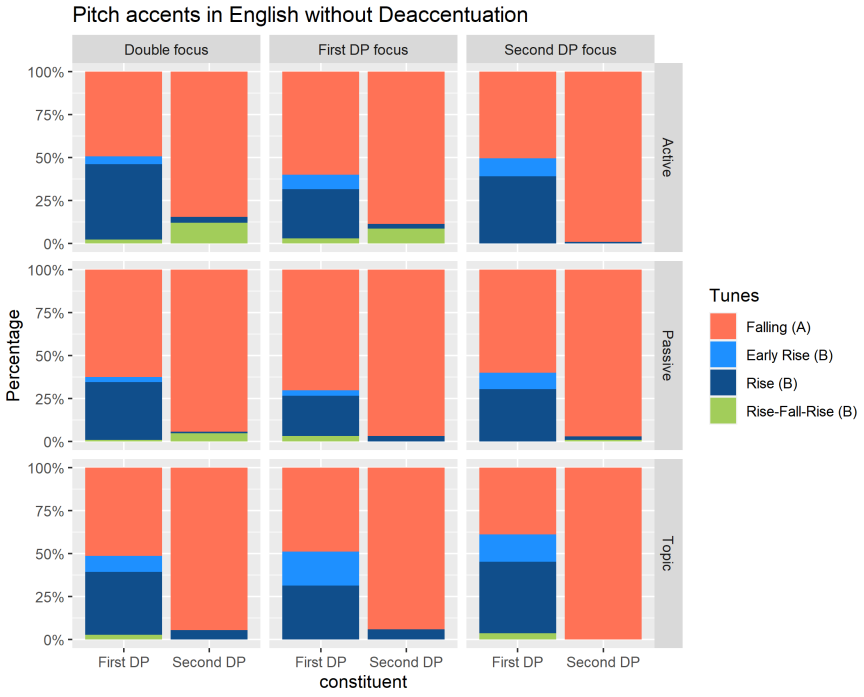


Fig. 2.15: Experiment 1.2: Pitch Accents without deaccentuation in English

Figure 2.16a shows the distribution of *A accents* across different conditions. As in Experiment 1.1, it becomes clear that overall more A accents were produced on the sec-

ond DP than on the first DP. A poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *A accents* was taken as independent variable and as dependent variables *focus*, *syntax* and *linear order* were included in the model. The final model performed significantly better than an intercept-only base line model (χ^2(5): 82.258, p = 0) and is a good, moderate fit (C: 0.7977297, Somers' Dxy: 0.5954594). The model's explanatory power is moderate (Nagelkerke's R2 = 0.16).[14] From the final model it becomes clear that only *linear order* has a significantly effect: More A accents were produced on the second DP than on the first DP (beta = 0.55, 95% CI [0.43, 0.68], p < .001. No other significant effects nor interaction effects were found.

The distribution of the Rise-Fall-Rise (B) accents is shown in Figure 2.16b. No clear pattern can be observed directly from this figure. A mixed-effects poisson regression model with random intercepts for speakers was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *Rise-Fall-Rise (B) accents* was taken as independent variable and as dependent variables *focus* and *syntax*. The final model performed significantly better than an intercept-only base line model (χ^2(4): 21.477, p = 0.0002546) and is a good, optimal fit (C: 0.9118041, Somers' Dxy: 0.8236082). The model's total explanatory power is substantial (conditional R2 = 0.35) and the part related to the fixed effects alone (marginal R2) is of 0.12.[15] From the final model it becomes clear that the effect of *syntax* is significant: More Rise-Fall-Rise (B) accents were produced in the active condition than the other conditions (see full model in Appendix B for exact estimates). *Focus* also had a significant effect: More Rise-Fall-Rise (B) accents were produced in the *Double DP focus* condition than in the *Second DP focus* condition (beta = -1.76, 95% CI [-2.83, -0.69], p = 0.001).

[14]The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = Double focus, is at -0.67 (95% CI [-0.82, -0.52], p < .001).

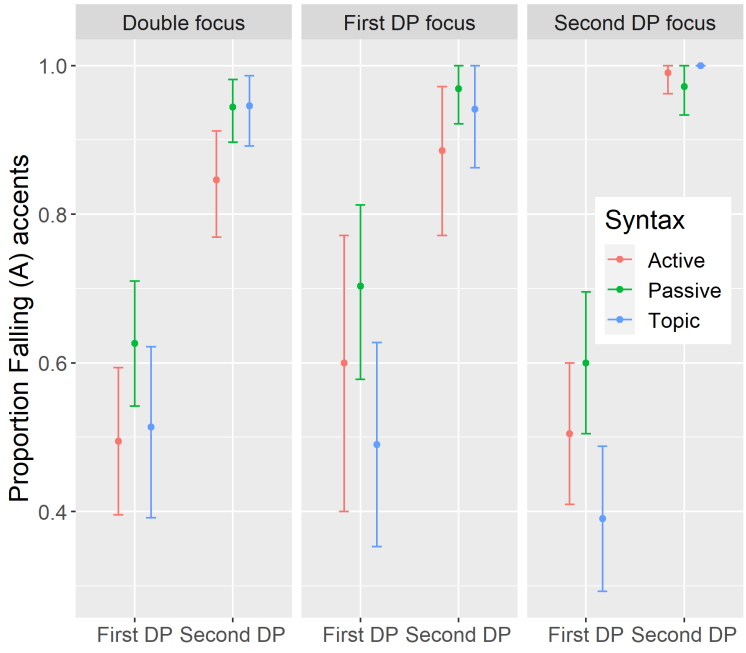
[15]The model's intercept, corresponding to syntax = Active and focus = Double focus, is at -3.48 (95% CI [-4.43, -2.53], p < .001).

The distribution of the Rising B accent can be seen in Figure 2.17a. From the figure it becomes clear that more rise accents were produced on the firs DP than on the second DP. A mixed-effects poisson regression model with random intercepts for speakers was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *R-B accents* was taken as independent variable and as dependent variables *focus*, *syntax* and *linear order* were included in the model. The final model performed significantly better than an intercept-only base line model (χ^2(5): 251.71, p = 0) and is a good, optimal fit (C: 0.8991235, Somers' Dxy: 0.7982469). The model's total explanatory power is substantial (conditional R2 = 0.53) and the part related to the fixed effects alone (marginal R2) is of 0.41.[16] From the final model it becomes clear that the effect of *linear order* is significant: Less R-B accents were produced on the second DP than on the first DP (beta = -2.69, 95% CI [-3.18, -2.21], p < .001).

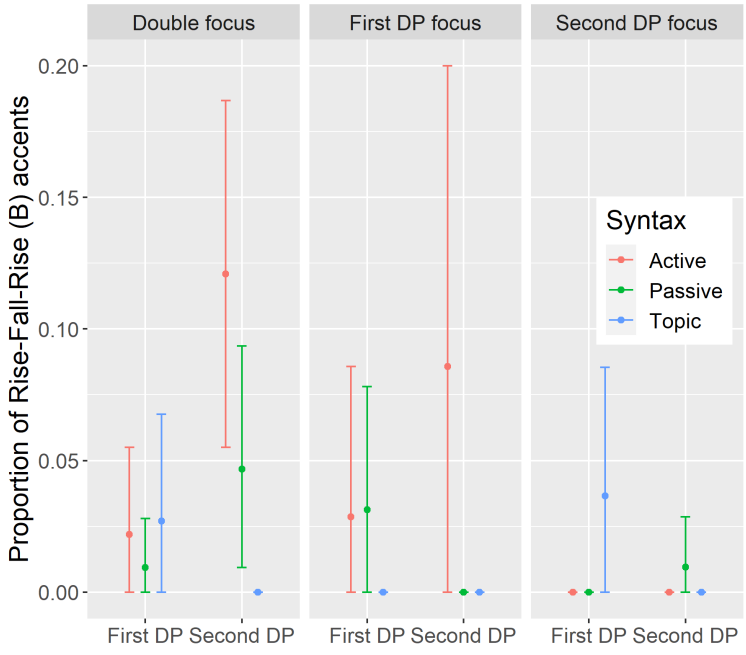
Finally, the distribution of the *Early Rises* across the different conditions is shown in Figure 2.17b. Again, more early rises were produced on the first DP than on the second, in fact no early rsies were produced on the second DP. In addition it seems that most early rises were produced in the topic word order and in terms of focus conditions most early rises were produced in the *First DP focus* condition. A similar a mixed-effects poisson regression to the one for rise (B) accents was performed. The final model performed significantly better than an intercept-only base line model (χ^2(5): 106.07, p = 0) and is a good, optimal fit (C: 0.9373336, Somers' Dxy: 0.8746671). The model's total explanatory power is substantial (conditional R2 = 0.96) and the part related to the fixed effects alone (marginal R2) is of 0.94.[17]. A significant effect was found for *syntax*: Most early rises were produced in the topic word order (beta = 0.65, 95% CI [0.04, 1.25], p = 0.036). It should also be noted that no statistical analysis could be performed on the linear order condition since there were zero productions on the second DP.

[16]The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = Double focus, is at -1.17 (95% CI [-1.55, -0.79], p < .001).

[17]The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = Double focus, is at -3.58 (95% CI [-4.44, -2.73], p < .001).

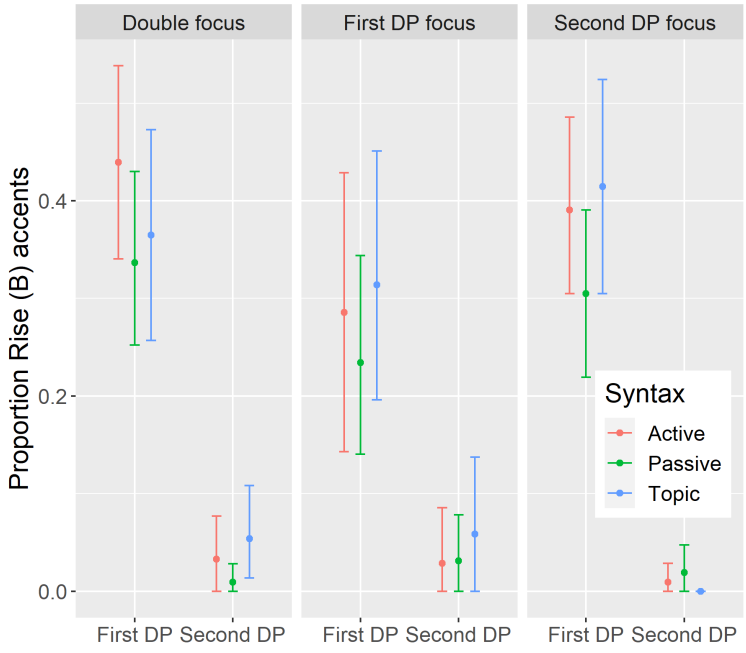


(a) A accents

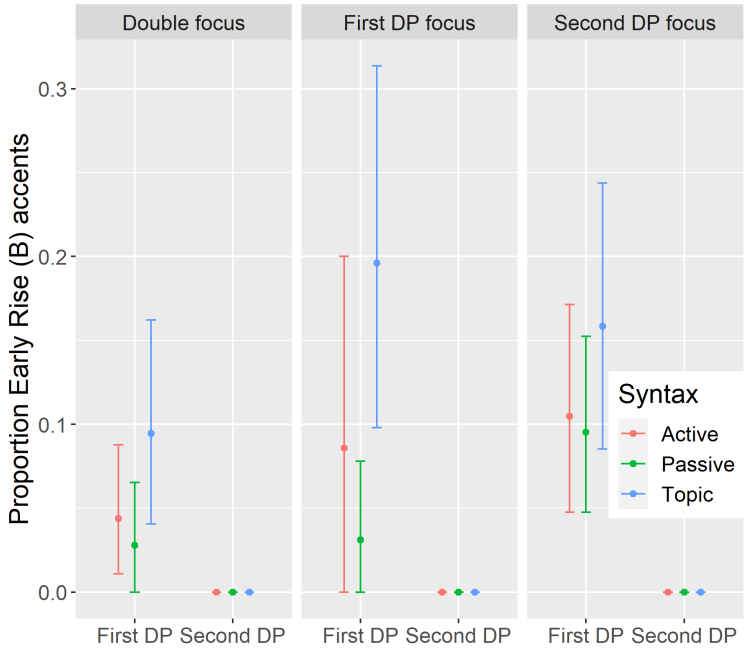


(b) Rise-Fall-Rise (B) accents

Fig. 2.16: Experiment 1.2: Distribution A and Rise-Fall-Rise (B) accents



(a) Rise (B) accents



(b) Early Rise (B) accents

Fig. 2.17: Experiment 1.2: Distribution Rise and Early Rise B accents

2.7 Conclusion & Discussion

Jackendoff (1972) and Büring (1997) hypothesized that contrastive topics are marked by B accents and foci by A accents. The experiments in this chapter set out to test this hypothesis by manipulating the focus contexts: The results clearly show us that the BA pattern does not convey whether the B-accented or the A-accented constituent functions as a contrastive topic in the utterance. In the *First DP focus* condition the first DP constitutes the focus and the second DP constitutes the topic, one would therefore expect more A accents on the first DP than on the second DP, and vice versa for the B accent. For the *Second DP focus* condition the opposite pattern was expected, i.e. a lot of B accents on the first DP and a lot of A accent on the second DP. Contrary to Jackendoff (1972), very few Rise-Fall-Rise (B) accents were produced overall and the focus conditions did not correlate with the type of pitch accent (i.e. not more AB patterns in the *First DP focus* condition and BA in the *Second DP focus* condition). Even when taking a less restrictive definition of the B accent, by including rises and early rises as B accents, the expected pattern by Jackendoff (1972) and Büring (1997) is still not observed.

Crucially the high proportion of B accents (be it RFR, Early Rise or Rise) on the first DP regardless of the focus condition goes against the predictions made by Jackendoff (1972) and Büring (1997). One can therefore conclude that the patterns proposed by Jackendoff (1972) are not confirmed by the production experiment: Different focus conditions did not switch around A and B accents, and thus B and A accents do not mark whether something is a topic or focus. Instead a strong effect of *linear order* for most pitch accents was reported which indicates that whether a pitch accent occurs on the first or second DP is strongly correlated with the type of pitch accent: There is a higher proportion of A accents on the second DP than on the first and the inverse is true for B accents. These results most strongly support the simple-focus analysis according to which there should be no significant difference in pitch accents across the different focus conditions instead the *lin-*

ear order determines the type of accent: B accents simply indicates that it is not the final accent whereas A accents do not come with such a restriction.

Like Experiment 1.1, Wagner (2012)'s nested account is not supported by the results of Experiment 1.2. According to this account there should be no BA patterns in the *First DP focus* since the nested focus analysis does not allow for contrastive topics to follow foci. However, from the results it became clear that many BA patterns were produced in the *First DP focus* condition.

Again, more evidence against Wagner (2012)'s nested foci account comes from the acceptability judgements. In general the topic sentences are judged least acceptable, followed by the passive ones, but crucially no difference across the different focus conditions can be observed at least for the active and topic cases (see Figure 2.18). In contrast to the other word orders, the passive sentences do seem to display the pattern expected by Wagner (2012), i.e. lower acceptability ratings for the *First DP Focus* condition compared to the other focus condition in sentences with both DPs being accented. It is not completely clear why this would be the case, but this pattern seems to have more to do with the passive syntax behaving differently from other word orders rather than an overall lower rating for the *First DP Focus* condition.

Let us now turn our attention to the main research question addressed in Chapter 1 about how intonational form and meaning are related to each other and whether intonational meaning is best analyzed compositionally or holistically. Contrary to claims made by Jackendoff (1972) and Büring (1997), this chapter does not support the claim that intonational meaning related to ABBA patterns is compositional and that there is a clear correspondence between a pitch accent and a specific meaning: The type of pitch accent is determined more by the order in which the pitch accents occur rather than a meaningful component in the sentence itself. There is no evidence supporting Jackendoff (1972) and Büring (1997)'s claim that pitch accents are morphological units that convey meaning. On the other hand, a holistic approach to the ABBA patterns is also not fully supported by the

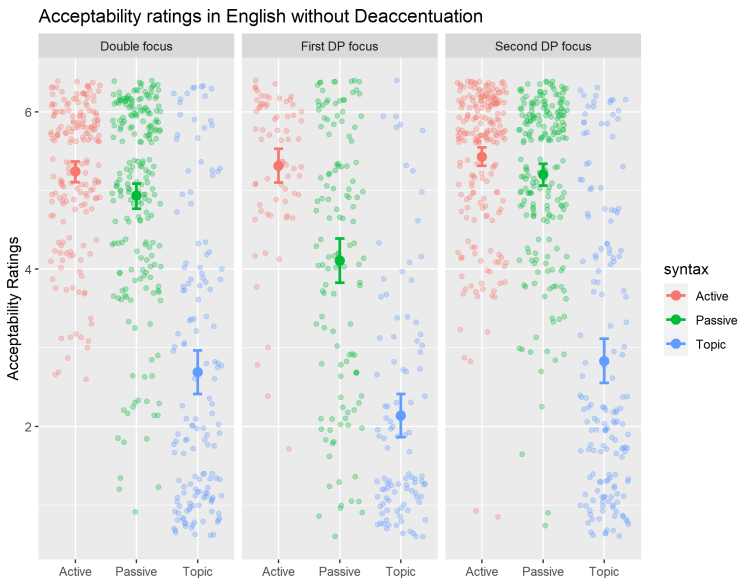


Fig. 2.18: Experiment 1.2: Acceptability Judgements

results, there is no clear meaningful component that can be attributed to the holistic tune. The only conclusion that can be drawn is that the results from this chapter do not support a compositional approach to intonational meaning as proposed by Jackendoff (1972) and Büring (1997). In the next chapter we will discuss an alternative analysis to the AB-BA patterns which might better capture the meaning component conveyed by these patterns.

In terms of accentuation an interesting observation can be made (see Figures 2.6 and 2.14): Deaccentuation of the second DP is expected given that only the first DP is focused in the *First DP focus* condition, however the opposite pattern would have been expected for the *Second DP focus*, i.e. deaccentuation on the first DP. There thus seems to be a general trend towards accenting the first DP regardless of whether it is focused or not. A possible explanation for this observation is provided by Calhoun (2012).

2.7.1 Theme vs. Rheme

Jackendoff (1972)'s claims have been put to the test by other studies as well like Calhoun (2007; 2012). Steedman (2000); Calhoun (2007) make different assumptions about focus semantics than Jackendoff (1972) or Rooth (1992), however. They position that in addi-

tion to a focus-background condition, there is a theme-rheme distinction. Calhoun (2012) follows Steedman (2000) by defining the theme to be "that part of the utterance which connects it to the rest of the discourse" and the rheme to be "the part of the utterance that advances the theme, or which is predicated" (Steedman 2000: p.655). Crucially, the notion of 'theme' arguably conflates the case of deaccented non-foci and accented non-foci. Example 39 illustrates the contexts used by Calhoun (2012) to evoke themes and rhemes.

(39) (Calhoun 2012: p.332)

Driver: If we do this, will the banana land on some money?

Slider: No the LOLLIPOP will land on some MONEY,
rheme *theme*
the BANANA will land on a MONSTER.
theme *rheme*

(40) Driver: If we do this, will the banana land on some money?

Slider: No the LOLLIPOP will land on some money.

Calhoun (2007; 2012) argues that there is an additional dimension that distinguishes these cases, which is the notion of contrast. Calhoun regards an object to be contrastive "if it contrasted with an equivalent object in the same utterance" (Calhoun 2012: p.332): In case of Example 39, *lollipop* would contrast with *banana* and *money* with *monster*. As already noticed by an anonymous reviewer, there is no guarantee that if the context provides a potential for a contrast, a speaker will actually encode the contrast. An objection was made that only rheme foci were truly contrastive in this experimental setup: Arguably, while *lollipop* and *monster* seem truly contrastive, the speaker can optionally make the choice to mark a contrast for *money* and *banana*. Both *money* and *banana* have been previously mentioned in the context and would therefore preferably not be accented. This becomes especially apparent when the last sentence is omitted, see Example 40. Intuitively, the way the first sentence of the Slider is likely to be pronounced in (39) seems very similar to how it would be pronounced in (40).

Regardless of whether a speaker chooses to contrast *money* with another element in the utterance, as Calhoun claims for Example 39, the fact still remains that this might clash with a preference to not accent this word since it is given in the context. I therefore agree with the anonymous reviewer in that only rheme foci seem truly contrastive whereas themes can optionally be contrasted with another item in the utterance. Consequently, I would regard what Calhoun calls "themes" to not evoke alternatives and depending on their realization they can be either accented or deaccented.

Calhoun (2012) argues that the theme-rheme distinction is marked by a shift in prominence, however as mentioned before Calhoun's definition of theme conflates the case of deaccented non-foci and accented non-foci. The idea that theme-rheme distinction and contrastiveness is reflected in the metrical structure of an utterance is an interesting hypothesis, but the evidence in Calhoun (2012) does not tease these dimensions apart. In fact, the contexts tested are similar to the contexts here in that they are compatible with deaccented and accented realizations of the non-focal constituent. The reported acoustic results are parallel to the results obtained here for the deaccented cases, in that there is a clear asymmetry such that non-focus in final position is much more reduced prosodically. This is compatible with there being prosodic subordination involved. As is, the results thus potentially conflate two cases that must be distinguished.

The claim that themes are metrically less prominent than rhemes is potentially different from the prediction of standard theories of focus like Rooth (1992): The prediction that theme accents are less prominent even when they are contrastive and not deaccented. This claim can be empirically tested, but arguably the data in the literature so far does not motivate the assumption of a prosodically conveyed theme-rheme distinction in addition to prosodic focus marking.

2.7.2 Prosodic subordination

Even though Calhoun (2012)'s implementation of the theme-rheme distinction is potentially problematic, the idea that pitch range scaling rather than pitch accent plays a role in encoding topics and foci, is worth exploring. Other researchers like Bishop (2013) believed that prosodic subordination was likely a more reliable cue to encoding topic-focus configurations than the final rise allegedly associated with topics. One could make a case that an prosodic subordination approach could also account for the results in this chapter. If one follows Calhoun (2012)'s idea that topics (or themes as she calls it) are lower in pitch range than foci (or rhemes) and one takes into account a general pitch declination over the course of an utterance, then the results found in this chapter could be partially explained: In the condition with focus on the first DP, the second DP which functions as the topic would be lower in pitch range, plus it would be extra low due to pitch declination and this could be perceived as what I have labeled "deaccented". For the *second DP focus* condition, the pitch peak on the first DP (i.e. the topic) is expected to be equally high or higher than the peak on the second DP (i.e. the focus), this would explain the higher proportion of perceived double accented patterns in this condition.

However the results for the *double focus* contexts are very similar to the results for the *second DP focus* condition, but unlike the *second DP focus* contexts, the double focus contexts do not explicitly give rise to contrastive topics. It is therefore unlikely that pitch range is a salient cue to encoding topic-focus configurations. Nonetheless, one could make the case that in double focus contexts the speaker can make a decision on whether or not to interpret either one of the constituents as a contrastive topic. In such a case, one might find the pitch range difference in the double focus condition to be different from the pitch range difference in the *second DP focus* condition. This still remains to be explored in future research.

Regardless of whether or not there is a prosodic subordination that encodes the topic-focus configurations, the main point made in this chapter still stands: B accent and A

accents do not mark for topics and foci respectively, thus invalidating Jackendoff's and Büring's claims about contrastive topics. The notion of B accent as being deaccented or lower in pitch range than A accents is not what Jackendoff or Büring had in mind when discussing the B and A accents.

2.8 Outlook

What about Jackendoff (1972)'s intuition of an independent and a dependent focus? One way to think about Jackendoff (1972)'s intuition of independent vs. dependent foci is that focus structure reflects the left-to-right building (i.e. incremental building) of compositional meaning. If JOHN is pronounced with a rising accent, this may convey that John–as opposed to someone else—did something. The following linguistic material fleshes out what John did. Such an incremental theory of focus interpretation can explain Jackendoff (1972)'s intuition, and also Wagner's intuition that 'contrastive topics' always come before foci.

The lack of word order preferences, however, does not really give evidence for Wagner's claims about a strict left-to-right ordering of foci. But maybe Jackendoff's contexts are simply not picky enough to detect this asymmetry. It certainly seems that Jackendoff's idea is very intuitive.

Appendix A: Number of Pitch Accents[18]

	First DP focus		Second DP focus	
	First DP	Second DP	First DP	Second DP
Active				
Falling (A)	14	44	28	125
Early Rise (B)	14	1	36	0
Rise (B)	17	0	0	4
Rise-Fall-Rise (B)	0	0	0	0
Passive				
Falling (A)	31	79	41	127
Early Rise (B)	18	1	28	0
Rise (B)	34	3	62	6
Rise-Fall-Rise (B)	0	0	2	0
Topic				
Falling (A)	13	46	23	100
Early Rise (B)	26	2	51	3
Rise (B)	9	2	33	4
Rise-Fall-Rise (B)	2	0	0	0

Table 2.2: Experiment 1.1: Number of contours in sentences without deaccentuation across different focus conditions and linear order

[18]**Go to the OSF-page to access all collected data. Click here.**

	Double focus		First DP focus		Second DP focus	
	First DP	Second DP	First DP	Second DP	First DP	Second DP
Active						
Falling (A)	45	77	21	31	53	104
Early Rise (B)	4	0	3	0	11	0
Rise (B)	40	3	10	1	41	1
Rise-Fall-Rise (B)	2	11	1	3	0	0
Passive						
Falling (A)	67	101	45	62	63	102
Early Rise (B)	3	0	2	0	10	0
Rise (B)	36	1	15	2	32	2
Rise-Fall-Rise (B)	1	5	2	0	0	1
Topic						
Falling (A)	38	70	25	48	32	82
Early Rise (B)	7	0	10	0	13	0
Rise (B)	27	4	16	3	34	0
Rise-Fall-Rise (B)	2	0	0	0	3	0

Table 2.3: Experiment 1.2: Number of contours in sentences without deaccentuation across different focus conditions and linear order

Contour	**Active**	**Passive**	**Topic**
A_A	119	175	103
A_RFR	12	4	0
A_Deaccented	105	58	86
A_Rise	3	6	1
RFR_A	2	3	5
RFR_RFR	1	1	0
RFR_Deaccented	0	0	3
Deaccented_A	2	2	8
Deaccented_Deaccented	0	2	2
EarlyRise_A	18	15	29
EarlyRise_Deaccented	0	1	12
Rise_A	89	83	72
Rise_RFR	1	1	0
Rise_Deaccented	2	2	30

Table 2.4: Number of patterns across sentences with different word orders

Appendix B: Statistical Models

	Falling (A) accent
(Intercept)	$-1.22\ (0.11)^{***}$
constituentSecond DP	$1.22\ (0.09)^{***}$
syntaxPassive	$0.04\ (0.09)$
syntaxTopic	$-0.05\ (0.10)$
focusSecond DP Focus	$-0.06\ (0.08)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2.5: Statistical Models Experiment 1.1: A accent

	Rise (B) accent
(Intercept)	$-1.37\ (0.23)^{***}$
constituentSecond DP	$-2.48\ (0.24)^{***}$
syntaxPassive	$0.10\ (0.14)$
syntaxTopic	$-0.34\ (0.18)$
focusSecond DP Focus	$0.34\ (0.15)^{*}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2.6: Statistical Models Experiment 1.1: R-B accent

	Early Rise (B)
(Intercept)	$-1.40\ (0.17)^{***}$
constituentSecond DP	$-3.11\ (0.36)^{***}$
syntaxPassive	$-0.30\ (0.20)$
syntaxTopic	$0.55\ (0.18)^{**}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2.7: Statistical Models Experiment 1.1: Early Rise accent

	Falling (A) accent
(Intercept)	-0.67 $(0.08)^{***}$
constituentSecond DP	0.55 $(0.06)^{***}$
syntaxPassive	0.10 (0.07)
syntaxTopic	-0.01 (0.08)
focusFirst DP focus	0.05 (0.08)
focusSecond DP focus	0.02 (0.07)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2.8: Statistical Models Experiment 1.2: A accents

	Rise-Fall-Rise (B) accent
(Intercept)	-3.48 $(0.49)^{***}$
syntaxPassive	-0.87 $(0.42)^{*}$
syntaxTopic	-1.14 $(0.52)^{*}$
focusFirst DP focus	-0.50 (0.48)
focusSecond DP focus	-1.76 $(0.55)^{**}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2.9: Statistical Models Experiment 1.2: Rise-Fall-Rise (B) accents

	Rise (B) accent
(Intercept)	-1.17 $(0.19)^{***}$
constituentSecond DP	-2.69 $(0.25)^{***}$
syntaxPassive	-0.22 (0.15)
syntaxTopic	0.03 (0.15)
focusFirst DP focus	-0.31 (0.18)
focusSecond DP focus	-0.07 (0.13)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2.10: Statistical Models Experiment 1.2: R-B accents

	Early Rise (B)
(Intercept)	-3.58 $(0.44)^{***}$
constituentSecond DP	-17.82 (109.16)
syntaxPassive	-0.38 (0.35)
syntaxTopic	0.65 $(0.31)^{*}$
focusFirst DP focus	0.63 (0.38)
focusSecond DP focus	0.78 $(0.32)^{*}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2.11: Statistical Models Experiment 1.2: Early Rise accent

Chapter 3

Possibly True Alternatives in English

Chapter 2 tested Jackendoff (1972) and Büring (1997)'s analysis of the BA and AB patterns in English. The results showed that B accents do not consistently mark contrastive topics and A accents do not consistently mark focus. Jackendoff (1972) and Büring (1997)'s claim that there is a direct mapping between a pitch accent and a specific meaning were not supported by the results. This chapter tests an alternative explanation to these BA and AB patterns in English proposed by Wagner (2012). According to this analysis the AB and BA patterns in English are two unrelated contours.

3.1 Alternative propositions

In the previous chapter we have seen that according to Jackendoff (1972) sentences with contrastive topics like the ones that are produced with a BA and AB patterns in English are best analyzed as a focus operator that abstracts over a tuple.[1] The presupposition of such a sentence with multiple foci is that there is some tuple $< x, y >$ such that the proposition derived by filling in those values is true. These variables are filled in an inherently asymmetric way: The value for one variable is freely chosen (the 'independent variable'), while the other variable is filled in second (the 'dependent variable'). The dependent variable

[1]Recall that Jackendoff (1972)'s B accent corresponds to a Rising-Falling-Rising accent

corresponds to the primary focus, while Jackendoff (1972) states that the way the independent focus is interpreted captures what it means to be a 'topic'. Crucially, the dependent and primary focus is realized by an A accent and the independent focus by a B accent.

Instead of having one focus operator, Büring (1997) assumes two distinct focus operators, namely F (for focus) and CT (for contrastive topic). Whichever element in the sentence receives Focus will be part of the set of propositions that are considered well-formed alternatives, called the *Focus value*. The CT operator abstracts over a set containing different Focus values (Topic Value), i.e. a set of sets of propositions. The topic value is a set inside of which the Focus values are nested. Important to note for Büring (1997) is that such constructions trigger an implicature based on the topic-semantic value, namely that there must be at least one element (a set of propositions, i.e., a question) that is still disputable in the context. Like Jackendoff (1972) foci are realized by an A accent and contrastive topics by a B accent.

Crucially, the CT and FOC are in a free distribution meaning the two operators can occur in any order. Jackendoff (1972) and Büring (1997) thus argue that both AB patterns and BA patterns are possible in English, depending on whichever element is the contrastive topic: In a way the BA and AB patterns are mirror images of each other.

Wagner (2012) does not view the BA and AB patterns as mirror images of each other, but considers them two distinct intonation patterns.[2] Example (41) was brought up as evidence against the idea of B and A accents as specific CT and FOC operators. If a B accent would consistently mark contrastive topics then the placement of a B accent should follow a change in syntax: Since in Example (41) *beans* is the contrastive topic, the B accent is expected to be equally acceptable in both (41a) and (41b). According to Wagner (2012) speakers strongly prefer not to accent the *beans*, since it is given in the context, and accenting it requires accommodating something that is not apparent in the discourse con-

[2]The B accents that Wagner (2012) has in mind correspond to Bolinger (1958)'s B accent, i.e. a rising accent.

text. The results from Chapter 2 support this observation, B accents did not consistently coincide with contrastive topics.

(41) (Wagner 2012: p. 22)

A: John ate the spinach.

B: And what about the beans? Who ate those?

a. B A
The BEANS, FRED ate.

b. A B
? FRED ate the BEANS.

Rather than linking AB and BA patterns to contrastive topics specifically, Wagner (2012) analyses these patterns and specifically the AB pattern (referred to as the Rise-Fall-Rise pattern) as operators over alternative sets, this is in line with Oshima (2005) and Constant (2012)'s approach to RFR patterns in English. Constant (2012) holds the view that RFR patterns in English introduce a conventional implicature that 'none of [the evoked] alternatives can safely be claimed'. Wagner (2012) on the other hand argues that the RFR patterns presupposes that an alternative is possibly true. The AB pattern is equated to the RFR pattern and would therefore carry the same semantics associated with this tune.

It should be noted that the meaning for the AB pattern in English closely resembles the semantics ascribed to the hat pattern. The hat pattern however, differs from the RFR pattern in that it presupposes that an alternative proposition is true, rather than just possibly true according to Wagner (2012). This chapter focuses solely on the intonation patterns in English, the following chapters will discuss hat patterns in German and Dutch in light of Wagner (2012)'s analysis.

Crucial to Wagner (2012)'s analysis of contrastive topic constructions in English is the fact that AB and BA patterns are not mere mirror images of each other, but instead constitute two distinct intonation patterns. Evidence that these two patterns are semantically different, comes from the fact that AB patterns are infelicitous as the last answer of a pair-list question (see Example 42b), but BA patterns are not (see 42a). This observation was

first noted by Krifka (1998), and later by Lee (2008) and Constant (2012). Wagner (2012) also notes that Example (42b) is not completely unacceptable, but the sense one gets from employing an AB pattern in the answer is that the speaker wants to insinuate something above and beyond what is conveyed by the pair-list answer itself.

(42) (Wagner 2012: p. 28)

A: Who kissed whom?

a. B A B A
B: ANNA kissed JOHN, and JIM kissed BERTA.

b. A B A B
?? B: ANNA kissed JOHN, and JIM kissed BERTA.

The AB and BA patterns do behave similarly when it comes to placing constraints on the discourse: Both require a salient contrastive alternative. In example 43 for instance, both the AB and BA patterns are somewhat infelicitous. According to Wagner (2012) this can be explained by the fact that the B accent requires a salient contrast to *John*.

(43) (Wagner 2012: p. 28)

A: It looks like someone already ate from the buffet!

a. A A
B: Yes, JOHN ate the BEANS

b. B A
#B: Yes, JOHN ate the BEANS

c. A B
#B: Yes, JOHN ate the BEANS

In contrast to the AB pattern, the BA pattern is acceptable even when all alternatives are resolved as true or false, see Example (44). The BA accent is felicitous here since there is a contextually provided contrast to the constituent produced with a B accent, in this case *Mary* is contrasted with *John*. According to Wagner (2012) the AB pattern is infelicitous because the only salient alternative (John broke up with Mary) is already resolved as false and would therefore be incompatible with the semantics associated with the RFR pattern.

Wagner (2012) also notes that the AA pattern is acceptable in all the dialogues presented in this section, including the ones involving contrastive topics. In contrast to the BA pattern which is argued to involve nested focus operators, the AA pattern is assumed to involve a single focus operator that binds to foci. It also seems that the AA pattern is the default choice for speakers when asked to produce these sentences without further instructions, this suggests that the use of a contrastive topic is not obligatory in these contexts (Wagner 2012). In contrast to Wagner (2012)'s nested foci analysis of BA patterns, the simple-focus analysis introduced in the previous chapter argues that such contours involve a single focus operator that binds to foci, much like the analysis proposed for AA patterns by Wagner (2012).

(44) (Wagner 2012: p. 29)

A: I heard that John and Mary split up. I wonder who dumped who.

a. A A
B: MARY broke up with JOHN

b. B A
B: MARY broke up with JOHN

c. A B
#B: MARY broke up with JOHN

3.2 Methodology

In order to test the claims made by Wagner (2012), an online production experiment was set up using jsPsych (de Leeuw 2015). Participants were presented, both in written form as well as auditorily, with a number of carefully manipulated questions, after which they were asked to record their answer and assess how well their answer sounded.

3.2.1 Participants

A total of 30 participants were recruited through Prolific, all of them were native speakers of North-American English. A compensation of 4.50 C$ was given which is based on an hourly rate of 13.50 C$ since the experiment itself took about 20 minutes.

3.2.2 Stimuli

Each stimulus consists of a short context, a question and an answer. The short contexts determine the number of possible alternatives. The questions consist of simple polar questions. For the answers there are five different conditions (see Figure 45): Answers with an *explicit true alternative* as illustrated in (45a), in this case *Valentina should call Frank* is a true alternative to *Ian should call Emma* and the affirmative *yes* ensures that the alternative is explicit. In the *no alternative correction* condition, the negative particle *no* excludes *Valentina should call Frank* as an alternative to *Ian should call Emma*, the latter instead is a correction to the former (45e). In the *implicit true alternative* condition, there is a true alternative to *Ian should call Emma* as indicated by the short context *Two phone calls should be made*, but it is not clear what the alternative is as indicated by the phrase *I don't know who should make the second phone call* (45b). The *open alternative* conditions makes no claims on whether there are any alternatives or not (45c). Finally, in the *no alternative question* condition, the participant is presented with a mutually exclusive choice between two alternative, therefore only one can be true and there is no alternative proposition true (45d).

(45) a. (Explicit True Alternative)

A: Two phone calls should be made. Should Valentina call Frank?

B: Yes. And I know what else: *Ian should call Emma*.

b. (Implicit True Alternative)

A: Two phone calls should be made. Should Valentina call Frank?

B: I don't know. But one thing is clear: *Ian should call Emma*. I don't know who should make the second phone call.

c. (Open Alternative)

A: I wonder if any phone calls should be made. Should Valentina call Frank?

B: I don't know. But one thing is clear: *Ian should call Emma*. Maybe that'll be good enough?

d. (No Alternative Question)

A: Should Ian call Emma or should Emma call Ian?

B: Well, that's pretty clear: *Ian should call Emma*. That should be more than enough.

e. (No Alternative Correction)

A: One phone call should be made. Should Valentina call Frank?

B: No. *Ian should call Emma*. That should be more than enough.

3.2.3 Procedure

Participants are directed to the experiment via an online link and the stimuli are presented to the participants in a latin-square design. Each trial starts off with a fixation point, next, the context and question are given in written form and finally on the next screen, the context and question are presented auditorily. Each question ends with a beep after which the participant is asked to record the answer as natural as possible. The trial finishes off with an acceptability input screen in which the participant has to indicate how natural their answer sounded, based on a scale from 1 to 8 where 1 means completely unnatural and 8 means completely natural.

3.3 Expectations

Although Jackendoff (1972)'s analysis to the BA and AB patterns is somewhat orthogonal to what is being tested in this chapter, the expectation would be that the BA and AB patterns are most appropriate in any context that allows for contrastive topics: Such contexts are the ones in which there are alternative propositions with which a contrast can be made. Based on this claim one expects the two contours to be least compatible with the conditions containing no alternative propositions and therefore the BA and AB patterns are expected to be most frequently produced in the *Explicit True Alternative*, *Implicit True Alternative* and *Open Alternative* conditions. Arguably, the *No Alternative Question* could be seen as containing a set of two contrastive alternative propositions and could be considered an acceptable context for the BA and AB patterns. The only prediction that can be made then based on Jackendoff (1972) is that no BA and AB patterns are expected in the *No Alternative Correction* condition. Moreover, potentially the BA pattern will be more frequently produced than the AB pattern assuming that in a sentence with two contrastive items the default is a BA pattern.

The *simple focus* analysis would make somewhat similar predictions to Jackendoff (1972), except that there are no AB accents, instead AA patterns are expected to be produced. In principle the expectation is that BA and AA patterns will be found in all conditions since in all conditions both DPs would be focus-marked, given that there is a double contrast in the answer to the polar question. Whether there are true alternatives or not is not important for the simple-focus analysis, what is essential is that the polar question evokes a double contrast in the answer.

Büring (1999) would expect the BA pattern (and AB pattern if a CT were to come last) to be most frequently produced in the the *Open Alternative* condition. Recall that in Büring (1997)'s analysis sentences containing contrastive topics come with a so-called disputability implicature, with disputable Büring (1997) means that there are informative but non-absurd answers that remain unresolved. Naturally, the most compatible condition to

this disputability implicature would be the *Open Alternative* condition. Furthermore, AA patterns are expected in all conditions.

Based on Wagner (2012) the expectation is that AB and BA patterns are felicitous in different contexts. As explained earlier, it is claimed that AB (or RFR in Wagner (2012)'s terminology) patterns in English presuppose that an alternative proposition is potentially true. The expectation would be that AB patterns are most frequently produced in the *Open Alternative* condition. The conditions with no contextually alternative propositions are predicted to be infelicitous with an AB pattern, the same is expected for the Explicit and Implicit True Alternative condition since the answer in this condition could be considered the last item in a pair-list answer which is incompatible with AB patterns according toWagner (2012).

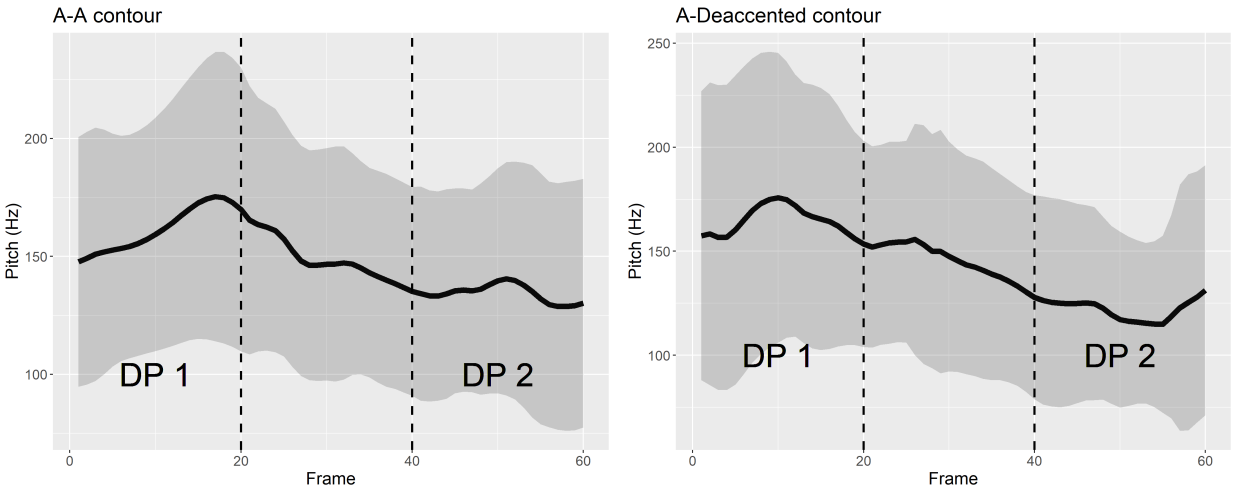
For the BA pattern the predictions are similar to the simple focus account: The BA requires contextually salient contrastive items, arguably all conditions contain such salient contrastive items. Additionally, in Wagner (2012)'s analysis BA patterns involve a nested focus structure, therefore the *No Alternative Correction* condition is expected to be less compatible with BA patterns since this condition does not allow for any alternative propositions. Finally, the AA pattern is expected to be most frequently produced in all conditions since it seems to function as a default contour in sentences with multiple foci. See Table 3.1 for an overview of the various predictions made by the different accounts.

	Explicit True	**Implicit True**	**Open Alter**	**Alter Question**	**Correction**
Jackendoff	AB/BA	AB/BA	AB/BA	AB/BA	
Büring	AA	AA	AB/BA/AA	AA	AA
Double Focus	BA/AA	BA/AA	BA/AA	BA/AA	BA/AA
Wagner	BA/AA	BA/AA	AB/AA	BA/AA	AA

Table 3.1: Expectations different focus accounts

3.4 Results: Mean Pitch Trajectory

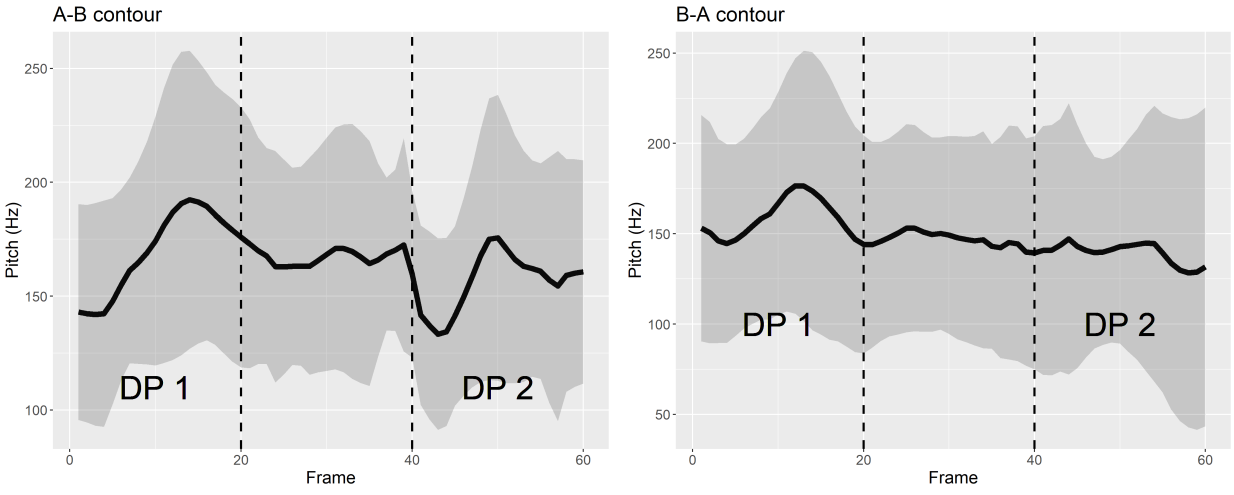
To give an idea of the actual productions, the mean pitch trajectories of some of the most frequent contours have been illustrated in this sections (see Figures 3.1 and 3.2). For the exact numbers see Table 3.2 in the Appendix. Besides the most common contours, the BA and AB patterns are provided as well since these were the ones most of interest to us (see Figure 2.13), however there are only few instances of these contours.



(a) AA pattern

(b) A-Deaccented contour

Fig. 3.1: Mean pitch trajectories, the black line indicates the mean pitch and the shaded area the standard deviation.



(a) A-RFR pattern

(b) RFR-A pattern

Fig. 3.2: Mean pitch trajectories.

3.5 Results

A total of 900 recordings were made (30 participants*30 trials = 900), three participant had to be excluded since one participant turned out to be non-native and for the other two no recordings were made. An additional 10 recordings had to be excluded due to poor sound quality, leaving us with a total of 800 recordings. Eventually, 627 out of 800 recordings contained no deaccentuations and were used for data analysis. For the results the recorded sentences were annotated by the author in terms of the pitch accents on the first DP and second DP. The following categories were distinguished for the annotation: Falling (A) accent, Rise-Fall-Rise (B) accent, Early Rise (B) accent, Rise (B) accent, Deaccented. For a phonological description of each pitch accent see the respective section in Chapter 2.

Figure 3.3 shows the proportions of pitch accents in percentage for each condition. Once again we are mainly interested in cases where both DPs are accented, i.e. without deaccentuation, therefore Figure 3.4 where all cases with deaccentuation have been excluded, is more informative: The exact numbers can be found in Tables 3.3 and 3.4 in Appendix A.

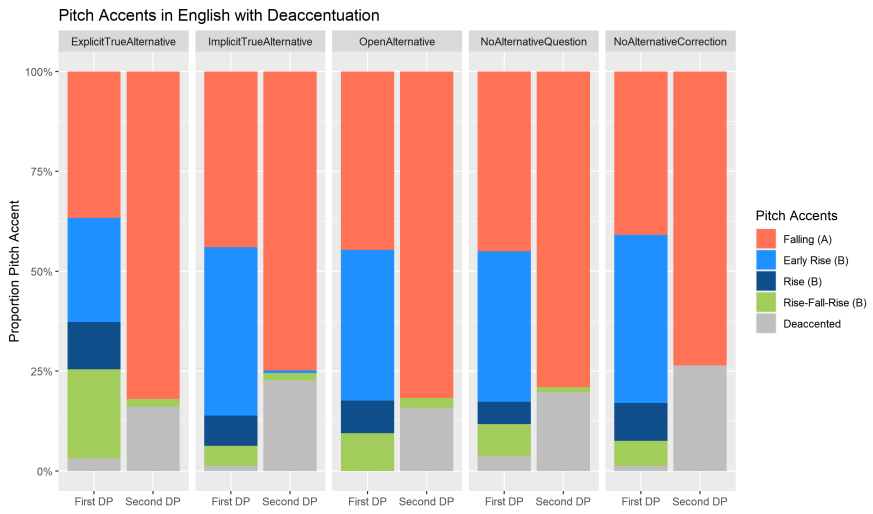


Fig. 3.3: Pitch Accents with Deaccentuation

Figure 3.5 gives a visualization of the distribution of the *Falling (A) accent* across different conditions. It becomes clear that there are more A accents on the second DP than on the

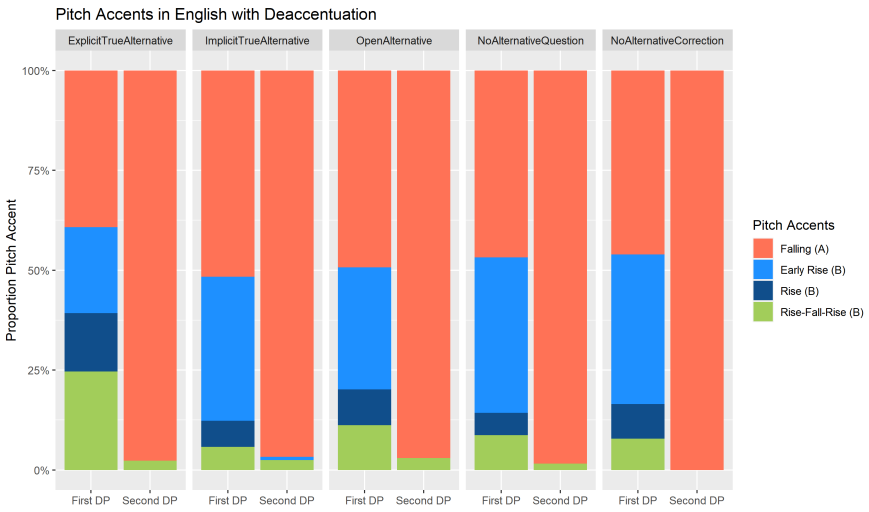


Fig. 3.4: Pitch Accents without Deaccentuation

first DP. To see whether there are any significant results a poisson regression was run.[3] As independent variable the number of *A accents* was taken and as dependent variables the *alternative* condition and *linear order* condition. The final model performed significantly better than an intercept-only base line model ($\chi^2(5)$: 117.69, p = 0) and is a good, optimal fit (C: 0.8322898, Somers' Dxy: 0.6645797). The model's explanatory power is moderate (Nagelkerke's R2 = 0.24).[4] From the final model it becomes clear that only *linear order* has a significantly effect: More A accents were produced on the second DP than on the first DP (beta = 0.74, 95% CI [0.61, 0.88], p < .001).

The distribution of the Rise-Fall-Rise accents across the different conditions is shown in Figure 3.6. Two observations can be made from this figure, first, most RFR accents are produced in contexts with explicit true alternatives and second, more RFR accents are produced on the first DP than on the second DP. A poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *RFR accents* was taken as independent variable and as dependent variables the *alternative* condition and *linear order* condition. The final model performed significantly

[3]See Appendix B for the full statistical models.
[4]The model's intercept, corresponding to constituent = First DP and assertion = ExplicitTrueAlternative, is at -0.82 (95% CI [-1.00, -0.65], p < .001).

better than an intercept-only base line model (χ^2(5): 71.704, p = 0) and is a good, optimal fit (C: 0.8770508, Somers' Dxy: 0.7541016). The model's explanatory power is moderate (Nagelkerke's R2 = 0.19).[5]. The model reports that only the *alternative* manipulation has a significant effect: There are significantly more RFR accents in the *ExplicitTrueAlternative* condition than in the other *alternative* conditions (see the full model in the appendix for the actual estimates).

The distribution of the Rise (B) accents across the different conditions is shown in Figure 3.7. Again, it seems that most rise accents are produced in the *Explicit True Alternative* condition and more rise accents have been produced on the first DP than on the second (in fact no rise accents are produced on the second DP). A poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *Rise (B) accents* was taken as independent variable and as dependent variables the *alternative* condition and *linear order* condition. The final model performed significantly better than an intercept-only base line model (χ^2(5): 84.286, p = 0) and is a good, not optimal fit (C: 0.81048778, Somers' Dxy: 0.6209754). The model's explanatory power is substantial (Nagelkerke's R2 = 0.27).[6]. The model reports that only the *alternative* manipulation has a significant effect: There are significantly more Rise (B) accents in the *ExplicitTrueAlternative* condition than in the *NoAlternativeQuestion* condition (beta = -0.97, 95% CI [-1.83, -0.10], p = 0.029). It should also be noted that no statistical analysis could be performed on the linear order condition since there were zero productions on the second DP.

Similar results were obtained for the Early Rise accent. The distribution of the Early Rises is shown in Figure 3.8. From this figure it becomes clear that most early rises are produced on the first DP and contrary to the other B accents less early rises are produced in the *ExplicitTrueAlternative* condition. Once again, a poisson regression model was fitted

[5]The model's intercept, corresponding to assertion = ExplicitTrueAlternative and constituent = First DP, is at -1.96 (95% CI [-2.57, -1.36], p < .001)

[6]The model's intercept, corresponding to assertion = ExplicitTrueAlternative and constituent = First DP, is at -1.92 (95% CI [-2.37, -1.47], p < .001)

to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *Early Rises* was taken as independent variable and as dependent variables the *alternative* condition and *linear order* condition. The final model performed significantly better than an intercept-only base line model ($\chi^2(5)$: 281.5, p = 0) and is a good, optimal fit (C: 0.8281447, Somers' Dxy: 0.6562894). The model's explanatory power is substantial (Nagelkerke's R2 = 0.45).[7]. The model reports that the *alternative* manipulation has a significant effect: There are significantly less Early Rise accents in the *ExplicitTrueAlternative* condition than in the other *alternative* conditions, except for the *OpenAlternative* condition which was not significantly different from the baseline (see the full model in the appendix for the actual estimates). A significant effect was also found for *linear order*: There are less early rises on the second DP than on the first DP (beta = -5.32, 95% CI [-8.19, -3.83], p < .001)

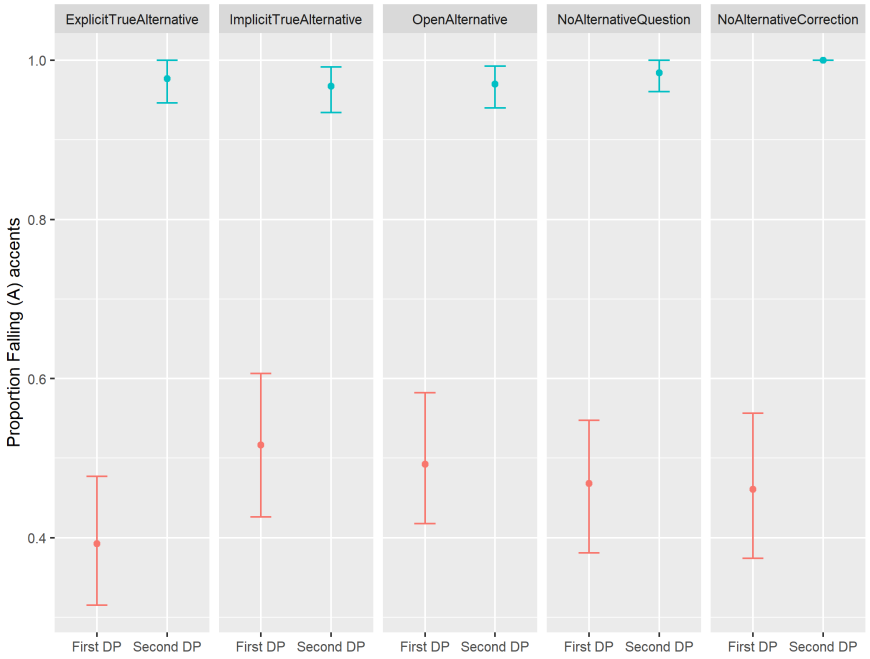


Fig. 3.5: Distribution of Falling (A) accents

[7]The model's intercept, corresponding to assertion = ExplicitTrueAlternative and constituent = First DP, is at -1.54 (95% CI [-1.94, -1.19], p < .001).

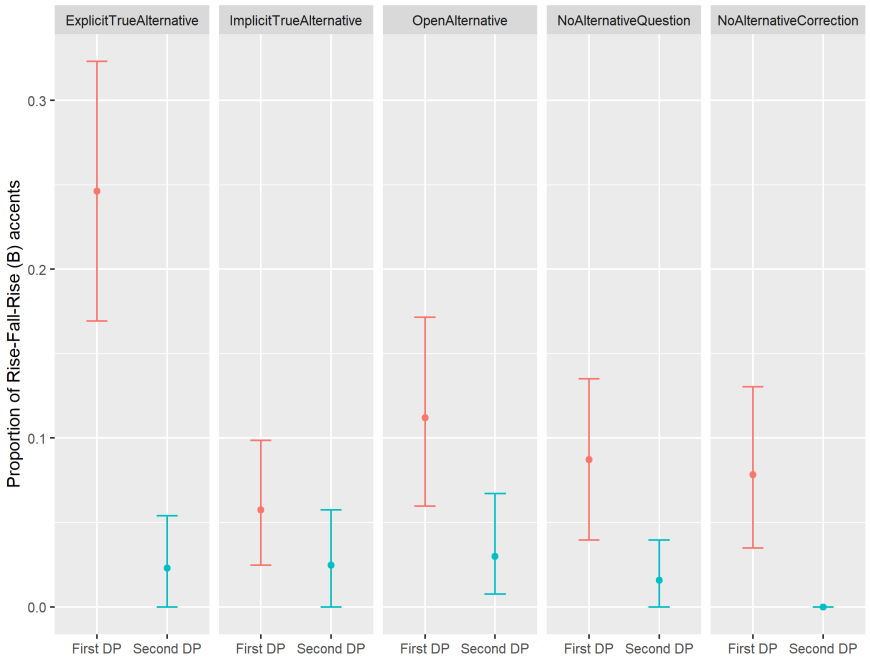


Fig. 3.6: Distribution of Rise-Fall-Rise (B) accents

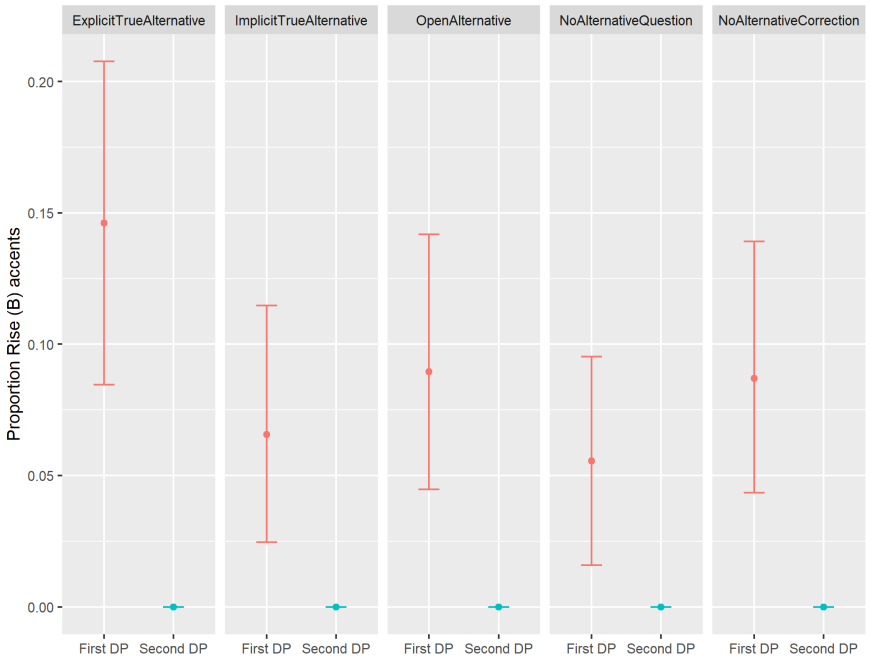


Fig. 3.7: Distribution of Rise (B) accents

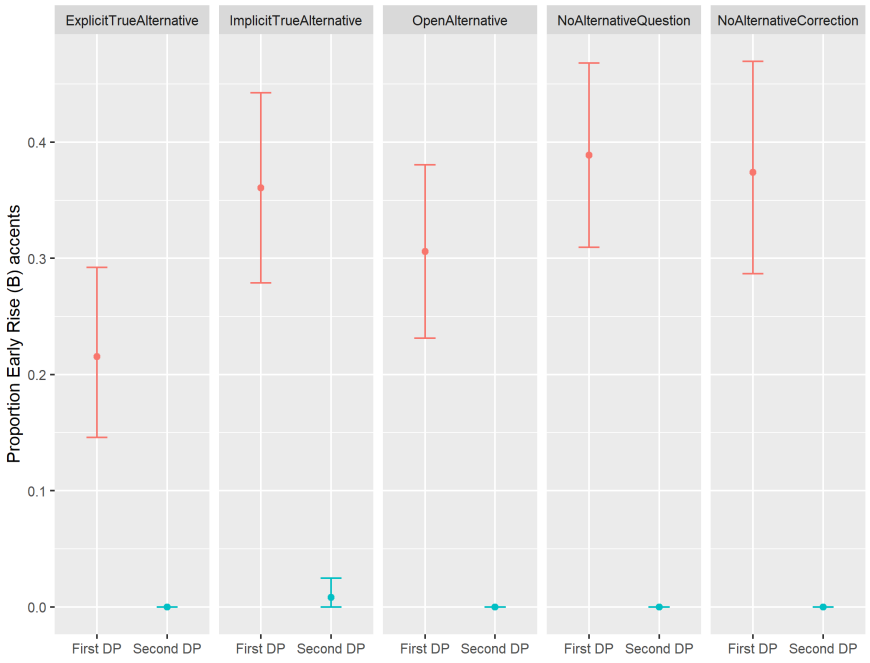


Fig. 3.8: Distribution of Early Rise (B) accents

3.6 Conclusion & Discussion

Interestingly none of the accounts seem to fully capture the results. Perhaps the simple-focus account comes closes to capturing the results. However, this analysis cannot account for the larger proportion of Jackendoff (1972)'s B accents and rising B accents in contexts with explicit true alternatives.

This brings us to Wagner (2008)'s analysis which was the one motivating this experiment. Wagner (2012) hypothesized that the AB pattern in English presupposes that an alternative proposition is potentially true and the expectation would be that this contour is most frequently produced in the *Open Alternative* condition. This claim is not supported by the results since there are very few B accents on the second DP.

Similar to the simple-focus account, Wagner (2012)'s nested focus analysis cannot account for the fact that BA patterns are be most prevalent in the *Explicit True Alternative* condition as reported by the results of this experiment. The BA pattern is not as frequently produced in the *No Alternative Question* and *Implicit True Alternative* conditions as

the *Explicit True Alternative* condition which is contrary to what was expected based on the idea that all these conditions contain a salient contrastive constituent in the context. This seems to suggest that BA patterns require not just a salient contrasting constituent, but an explicit true alternative. Interestingly, this outcome is what would be expected for the hat pattern in German and Dutch by Wagner (2012).

Returning to our main research question of whether intonational meaning is best analyzed holistically or compositionally, it can be concluded that for the ABBA patterns, a holistic approach is more supported by the data than a compositional approach.

Recall that in Chapter 2 the experiment was set up in such a way that the pragmatic/semantic function of the first and second DP in the sentence was determined by the question being asked, i.e. whether the object was being questioned (*What did Fred eat?*) or the subject (*Who ate the beans?*) and which constituent served as the topic through these *What about...?* questions. Such an experimental setup allowed us to determine whether specific meanings could be linked to specific contours: No such evidence was found, there was no strict correspondence between a pitch accent and a specific meaning component. The contexts in this experiment all pose the same polar question across the different conditions, in Example 45 the polar question was *Should Valentina call Frank*, hence not evoking a specific meaningful distinction on either the first or second DP. Instead the current experiment tested whether the existence of alternative propositions had an affect on the intonation contour being produced. The results of this experiment showed more (RFR and Rising) B accents on the first DP in contexts with explicit true alternatives even though the semantics of the first DP was not specifically manipulated.

All in all, one can conclude that the current experiment is in line with the results from the experiment in Chapter 2 in that neither of the experiments seem to support the idea that specific contours (that make up a tune) convey specific meanings. There does seem to be a correlation between a holistic pattern and an overall sentence meaning: The BA pattern is associated with contexts in which there is an explicit true alternative. In conclu-

sion, the results from chapters 2 and 3 show that AB-BA patterns do not convey whether the B-accented or the A-accented functions as a contrastive topic in the utterance, it simply conveys that two constituents evoke alternatives. Moreover, non-final accents can be optionally realized with a rise (i.e. a B accent).

3.6.1 AB patterns and Incomplete Answers

The results from this chapter reported very few AB patterns and so it can be concluded that the contexts used in this chapter do not capture what is expressed by AB patterns: The hypothesis that AB pattern in English presupposes that an alternative is potentially true, was not supported by the results. Naturally the question arises whether there is a context in which such an intonation pattern is more appropriate.

In fact there is such a context in which the AB pattern (or Rise-Fall-Rise contour as it is called in those studies) is most frequently produced: Both Wagner et al. (2013) and Goodhue et al. (2016) found the AB pattern to be produced about 65% of the time in contexts containing incomplete answers. An example of such an incomplete answer is provided in Example 46 below.

(46) (Goodhue et al. 2016: p.316) **Incomplete Response**

A: I don't feel like going to this party tonight, I have the feeling I might not like any of the people there.

[You know your friend John is attending the party, and you know Emma knows and likes him, but you're not sure whether she'll like anyone else, and your answer should reflect that.]

B: You like John

What this example shows is that simply requiring a context with open alternatives as was hypothesized in this chapter, is not enough to evoke AB patterns. Instead this

intonation pattern requires that there is something above and beyond the answer that is true, one cannot simply state that an alternative is potentially true: A context with open alternatives leaves it unsaid on whether there is something above and beyond the answer that is true. It seems that a stronger implicature is needed than originally hypothesized. This implicature that AB patterns express that something is true above and beyond the answer is very similar to Büring's disputability implicature which states that there is an alternative in the topic value that still remains to be resolved. However, as we have seen from the previous chapter, Büring's semantic representation of contrastive topics is not supported by the results.

Something that has not been mentioned but could give us a clue as to the semantics of the AB pattern is that Wagner (2012) actually argues that the RFR contour (or AB pattern) conveys that an alternative speech act is possibly true rather than an alternative propositions: "One way to capture the distribution of the RFR is to posit that the RFR takes an assertion rather than a proposition as its argument. This would also explain why it cannot occur in questions (they are not assertions but sets of propositions), and also why it cannot occur in complement clauses (they are not assertions but propositions). The meaning of the RFR then has to be characterized based on the applicability of an alternative assertive speech act rather than the possible truth of an alternative proposition (where S is a speech act, and DS is the set of all entities of the type that assertive speech acts have, and assuming that a notion of entailment between speech acts is defined)" (Wagner 2012: p.26). This way of looking at it might also help explain why it seems that speakers use the AB pattern when they do (in for instance incomplete answers), they use the AB pattern when they want to imply another speech act and when they want the interlocutor to draw an additional inference.

To conclude, although we know that AB patterns have something to do with incomplete answers, it still remains unclear how to formally explain the semantics behind such

incomplete answers: All one can conclude is that having contexts with open alternative is not sufficient to evoke AB patterns.

Appendix A: Number of Pitch Accents[8]

Contour	Explicit	Implicit	Open	NoAltQuestion	NoAltCorrection
A accent_A accent	48	61	65	59	53
A accent_RFR accent	3	2	1	0	0
Early Rise_A accent	28	43	40	47	43
Early Rise_Early Rise	0	1	0	0	0
Early Rise_RFR accent	0	0	1	2	0
RFR accent_A accent	32	6	13	11	9
RFR accent_RFR accent	0	1	2	0	0
Rise accent_A accent	19	8	12	7	10

Table 3.2: Number of contours across sentences with different word orders

	Explicit		Implicit		Open	
	First DP	Second DP	First DP	Second DP	First DP	Second DP
Falling (A) accent	51	128	63	118	66	130
Early-Rise	28	0	44	1	41	0
Rise (B)	19	0	8	0	12	0
Rise-Fall-Rise (B)	32	3	7	3	15	4

Table 3.3: Number of contours across different assertion conditions

	NoAltQuestion		NoAltCorrection	
	First DP	Second DP	First DP	Second DP
Falling (A) accent	59	124	53	115
Early-Rise (B)	49	0	43	0
Rise (B)	7	0	10	0
Rise-Fall-Rise (B)	11	2	9	0

Table 3.4: Number of contours across different assertion conditions

[8]**Go to the OSF-page to access all collected data. Click here.**

Appendix B: Statistical Models

	Falling (A) accent
(Intercept)	$-0.82\ (0.09)^{***}$
constituentSecond DP	$0.74\ (0.07)^{***}$
assertionImplicitTrueAlternative	$0.08\ (0.11)$
assertionOpenAlternative	$0.07\ (0.10)$
assertionNoAlternativeQuestion	$0.06\ (0.11)$
assertionNoAlternativeCorrection	$0.06\ (0.11)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.5: Statistical Models: A accent

	Rise-Fall-Rise (B) accent
(Intercept)	$-1.96\ (0.31)^{***}$
constituentSecond DP	$-1.82\ (0.31)^{***}$
assertionImplicitTrueAlternative	$-1.19\ (0.36)^{***}$
assertionOpenAlternative	$-0.78\ (0.29)^{**}$
assertionNoAlternativeQuestion	$-1.03\ (0.33)^{**}$
assertionNoAlternativeCorrection	$-1.31\ (0.38)^{***}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.6: Statistical Models: RFRise (B) accent

	Rise (B) accent
(Intercept)	$-1.92\ (0.23)^{***}$
constituentSecond DP	$-18.86\ (1011.20)$
assertionImplicitTrueAlternative	$-0.80\ (0.42)$
assertionOpenAlternative	$-0.49\ (0.37)$
assertionNoAlternativeQuestion	$-0.97\ (0.44)^{*}$
assertionNoAlternativeCorrection	$-0.52\ (0.39)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.7: Statistical Models Experiment Ch3 :Rise (B) accent

	Early Rise (B) accent
(Intercept)	$-1.54\ (0.19)^{***}$
constituentSecond DP	$-5.32\ (1.00)^{***}$
assertionImplicitTrueAlternative	$0.54\ (0.24)^{*}$
assertionOpenAlternative	$0.35\ (0.25)$
assertionNoAlternativeQuestion	$0.59\ (0.24)^{*}$
assertionNoAlternativeCorrection	$0.55\ (0.24)^{*}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.8: Statistical Models: Early Rise

Part II: Hat Pattern

Chapter 4

Two nuclear falling f0 patterns

In the previous chapters we talked about BA and AB patterns in English and their semantics according to Jackendoff (1972). In this part of the dissertation an intonational pattern closely related to the BA and AB patterns will be introduced, namely the hat pattern in German and Dutch. Before turning our attention to the hat pattern specifically, we will digress a bit from the main topic and discuss two accents in German and Dutch that make up a hat pattern and that play a crucial role in distinguishing the different phonological types of hat patterns. Broadly speaking, the hat pattern consists of a rising gesture, followed by a high plateau and ends with a falling gesture, thus forming an apparent hat-shape in the pitch trajectory. The falling gesture of the hat pattern is said to be crucial to the meaning conveyed by the hat pattern and different falling gestures are said to convey different meanings (e.g. according to Wagner (2012) it is only hat patterns with an early peak that convey that an alternative proposition must be true). In terms of the phonological representation of these falling gestures, there is no consensus. The current chapter tackles this issue and makes an attempt at exploring the phonological representations of the different falling pitch gestures.

In Dutch there are two distinct nuclear falling f0 patterns (Gussenhoven 1991; Rietveld and Gussenhoven 1995; Caspers 2000; Braun and Tagliapietra 2010): the *'early'* peak

accent and the *'late'* peak accent. Similar accents have also been described for German (Grice et al. 2006; Rathcke and Harrington 2006). Roughly speaking, the label 'early' is used to indicate that the f0-peak occurs before (or early in) the accented vowel whereas the label 'late' indicates that the f0-peak occurs (late) in the accented syllable. There is no consensus on the exact phonological representation of these two accents, but as the names already suggest there seems to be a difference in peak alignment between the two accents. Especially the representation of the early peak accent is debated. This chapter discusses the different analyses for the early peak accent and tests whether first of all the early and late peak accent are phonetically different from each other and secondly which phonological representation best fit the phonetic results.

On the semantic side, Caspers (2000) claims that an early peak accent in Dutch is not compatible with new information whereas the late peak accent is. More specifically, according to Caspers (2000) the early peak accent is compatible with expected information whereas the late peak accent doesn't require any expectations. This semantic difference will be crucial for the manipulations in the current experiment.

4.1 Downstepped H tone

In the Transcription of Dutch Intonation (ToDI) framework, some have analyzed these two falling f0-patterns as a downstepped H starred tone with a trailing L tone (!H*L) for the *early peak* accent, and an H starred tone with a trailing L tone (H*L) for the *late peak* accent (Gussenhoven 1991; Rietveld and Gussenhoven 1995; Caspers 1997; 1999; 2000; Braun and Tagliapietra 2010). Downstep means that in a sequence of multiple H tones every subsequent H tone is slightly lower in pitch than the previous H tone (Ladd 2008). This representation suggests that the difference between the two accents is one of peak height (rather than peak alignment).

Except for a schematic representation (see Figure 4.1) Caspers (1999) provides little phonetic detail of what these different accentual tones would sound like. The phonological representation in ToDI seems to suggest that both accentual falls have a trailing L tone but only the early peak has a downstepped H tone. A downstepped H implies there is a previous triggering tone. For Caspers (1999) this previous tone is a high initial boundary tone. As the name already suggests, a downstepped H tone predicts a lower scaling of the beginning of the fall. In the cases illustrated in Figure 4.1 this means relative lower both to an initial high boundary tone (syntagmatically) and the non-downstepped version of the late accent lending fall (paradigmatically). What becomes apparent both from the IPO and ToDI analyses is that the early and late peak accents are phonologically similar, implying that they would have a more or less similar pitch trajectory: Whereas the IPO transcription seems to suggest an alignment difference (i.e. a difference on the horizontal temporal dimension) of the same fall, the ToDI system seems to suggest a qualitative difference in pitch level (i.e. a difference on the vertical pitch dimension) in which the early peak has a downstepped (lowered) high tone. No further clarification is given of the exact phonetic difference between these two accents, and both transcriptions are taken to be equivalent even though they differ along various dimensions (alignment vs. scaling).

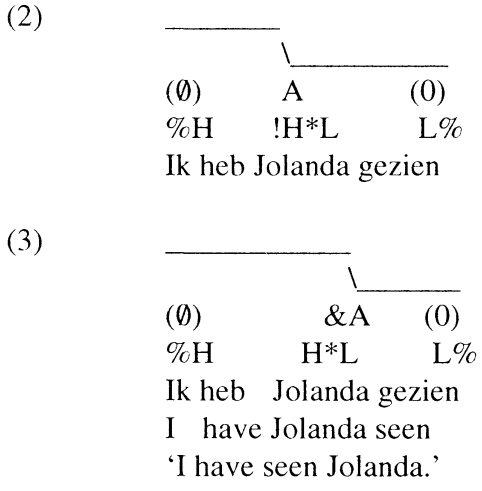


Fig. 4.1: Casper (2000) p.28: schematic representation of the early and late peak in Dutch.

Representing the early peak accent with a downstepped high tone implies that there must be a previous high tone triggering the downstep of the second H tone. There are three ways of meeting this requirement. If there is more than one pitch accent in the same intonational phrase (IP), then this high tone could belong to a preceding high pitch accent through spreading. If however the downstepped high tone is the only pitch accent within the IP then downstep could take place from a high initial boundary tone: This is the assumption Caspers (1997; 1999; 2000) make by representing the early peak accent as a combination of a high initial boundary tone followed by a downstepped high tone and a low trailing tone (%H !H*L L%).

It should be noted that if the difference between the two accents would be only downstep, as Caspers (1997; 1999; 2000) seems to suggest, then there should not be a difference in meaning: Downstep is a phonetic categorical process by which the scaling of a H tone is lowered, and where the identity of the tone remains constant. If there is a difference in meaning, we are dealing with two phonologically distinct accents, not with one that is sometimes downstepped (due to context), and sometimes not. However, if the meaning resides in the accent and not in the accent, one could argue that the downstepped version occurs in a accent that contains initial H tones, and the non down-stepped version occurs in a accent with a low tone in initial position, hence no downstep. [1]

A third possibility would be to posit a high leading tone in the syllable preceding the accented syllable. Gussenhoven et al. (2005) argues that the early peak accent should be transcribed as a high leading tone followed by a downstepped high tonal target (H!H*). [2]

The variety of possible analyses begs the question of whether the high tone preceding the downstepped pitch accent is a high boundary tone aligned with the left edge of the

[1]Clearly Caspers does not regard downstep as a phonetic process (see Figure 4.1) but a way to represent a certain intermediate phonological category that is neither H nor L. A "M(id)" tone would presumably be more appropriate. However since this introduces a whole paradigm shift in ToBI which only acknowledges two fundamental tones (i.e. H and L), the author may have chosen to "incorrectly" represent the early peak accent as a downstepped H tone.

[2]Interestingly, not only is this different from his earlier analysis, it also runs against his own ToDI system which strictly speaking does not allow for leading tones.

utterance, a high leading tone aligned early in or slightly before the accented syllable, or a spreading tone originating from a previous H accent. Regardless of whether in the early peak accent the nuclear tone is preceded by a high leading tone or a high initial boundary tone, both preceding tones would be obligatorily part of the accent in order to motivate a downstepped nuclear tone (that gives rise to a meaningful difference).

4.2 Early and Late Peaks in German

In the literature on German different falls have been proposed, besides a late peak (H*) and early peak (HL*), Grice et al. (2006) argue there is an intermediate fall (transcribed as H*!H). According to Grice et al. (2006) the semantic contrast that is evoked by the latter two tonal distinctions in German can be summarized as that between general or polite statements (H+!H*) and resolute or soothing assertions (H+L*).

Similar to Grice et al. (2006), Rathcke and Harrington (2006) suggest that in German the difference between the early and late peak is one of scaling as well, however they question the phonological status of the intermediate fall. In an experiment they scaled the peak of the accentual fall by lowering the peak stepwise (see figure 4.2 and table 4.3 below). The goal of the experiment was to distinguish an intermediate peak from an early peak as separate phonological categories. In an experiment two different contexts where created based on the description above, participants were presented with auditory stimuli of the different accentual falls and then they were asked which of the two contexts was most appropriate or if neither was appropriate.

The results of the semantic congruity test showed that there was a significant difference when going from step 2 to 3, but any subsequent steps were not significant. Based on this result the conclusion was drawn that participants phonologically distinguish H* from other falling accents (i.e. H+L* and H+!H*), but that there was no need to posit an

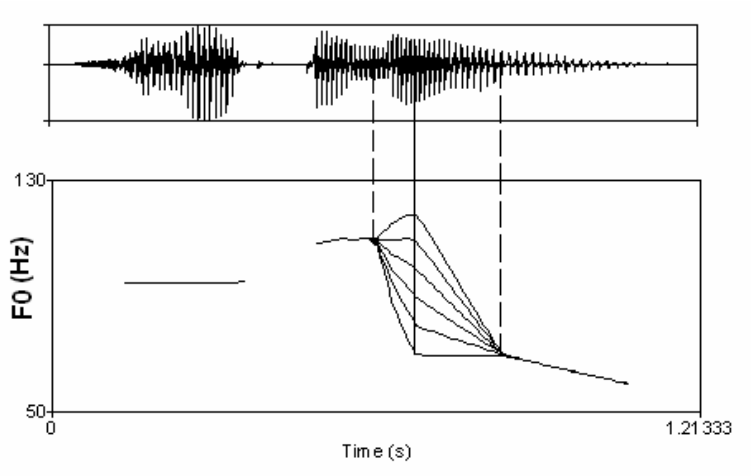


Fig. 4.2: Rathcke & harrington (2006) p.2: Synchronized time-waveform and the six tokens of the synthetic continuum. The dashed lines mark the first fixed reference point preceding the accented vowel and the second reference point at the vowel offset, respectively. The solid vertical line shows the manipulation point in the accented vowel.

Table 1: *F0-values in the accented vowel and the pitch accent categories to which the F0-manipulations are assumed to correspond (column 3)*

stimulus number	*F0 at the vowel (Hz)*	*pitch accent (target at the vowel)*
1	120	H* (high target)
2	110	
3	100	H+!H* (mid target)
4	90	
5	80	H+L* (low target)
6	70	

Fig. 4.3: Rathcke & harrington (2006) p.2: Overview of Stimuli

intermediate fall category. In other words there is evidence for a binary distinction but not a distinction between three categories.

Besides the semantic congruity test, a perception experiment in the form of an ABX task was set up as well as an production experiment in the form of an imitation task. The imitation results supported the semantic congruity test in that step 2 and 3 were significantly different in production. In addition, a significant effect was also found between step 1 and 2, however, Rathcke and Harrington (2006) argue that participants were hearing a gradient difference in prominence rather than a difference in phonological tonal cat-

egories. Unlike in the semantic congruity test from which it was concluded that there is a binary distinction, the discrimination task showed that participants were able to perceive the six differently scaled falls. Given these results, it was concluded that the difference between H+!H* and H+L* is a phonetic one and only H* is phonologically different from the other two pitch targets. In other words, phonologically speaking there is only a late peak accent (represented as H*) and an early peak accent (represented as H+L*).[3]

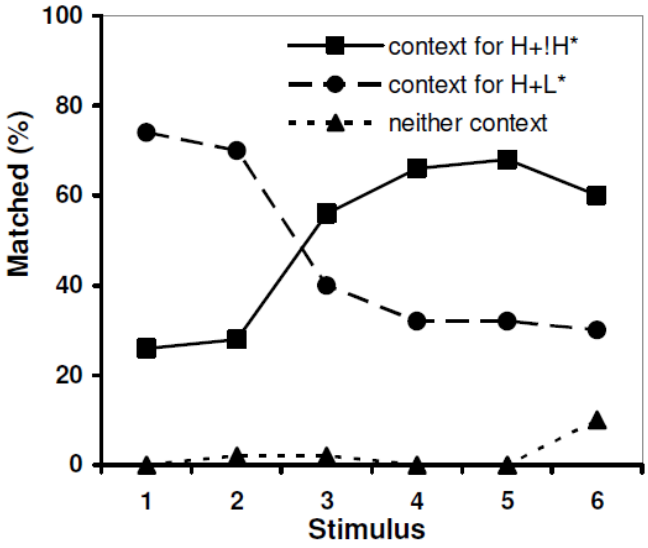


Fig. 4.4: Rathcke & Harrington (2006) p.3: Results of semantic congruity tests: percentage of 'matching' judgements for 6 stimuli depending on the context (n = 10).

Having established that there is a two-way phonological distinction in terms of falling patterns in German[4], Rathcke & Harrington continued researching the exact phonetic and phonological difference between the two accents (Rathcke and Harrington 2007).

In an imitation experiment they looked into the scaling and alignment of the H and L tones in the two different falling patterns in German. According to GToBI, late peak accents are best analysed as monotonal H* followed by a L boundary tone, early peak accent in contrast consist of a bitonal H+L* pitch accent followed by a L boundary tone. The

[3]Rathcke and Harrington (2006) remarks that the early peak accent could have been represented with H+!H* as well, and the choice to represent it with H+L* was more or less arbitrary. Crucially however, there is no evidence to consider H+!H* and H+L* two phonologically distinct categories

[4]It should be noted that this binary distinction was only found for the semantic congruity test and partially for the imitation task.

phonetic difference in peak alignment between late and early peaks in GToBI is analysed as a difference between unassociated and associated H-tones with L as a starred tone in an early peak and as a boundary tone in a late peak.

Table 1. *Overview of test items selected as visual stimuli for the imitation task.*

Syllable count after the nucleus	Syllable structure	
	CVC	CV:
none	*Linn*	*Lie*
one	*Linner*	*Liener*
two	*Linnerer*	*Lienerer*

Fig. 4.5: Rathcke & Harrington (2007) p.2: Overview of Stimuli

Based on the results, several conclusions were drawn. First, early and late peaks mainly differ in the alignment of H tonal target: it is aligned in the first half of the syllable for an early peak and in the second half for a late peak. In both cases, a leftward displacement of high f0-targets occurs under time pressure caused by decreasing the amount of segmental material following the accented syllable. Secondly, in terms of the scaling of the L tone, L tones are lower in the early peak than in the late peak accent.

The alignment of low targets is unstable both in early and in late peaks. Furthermore, the alignments of the low targets in both pitch accents do not differ from each other, which is surprising if we assume that an L-tone has a different status depending on whether it is an associated tone in a bitonal pitch accent or a boundary tone preceding a phrase edge. However, the interval between the right-hand phrasal edge and the L-target is not fixed, the alignment of the low tone is more variable: it is not temporally anchored to the phrase boundary nor is it systematically different for the two pitch accents.

On the other hand, the finding that the L-scaling is lower in early peaks as well as perceptual investigations (Niebuhr 2003; Rathcke and Harrington 2010), are consistent with an analysis of the early peaks as HL*. As has been shown in Niebuhr (2003), the perceptual difference between late and early peaks depends predominantly on the perception of

the pitch fall: relatively steep falls through the accented syllable (as a result of which the low target is reached faster than in late peaks) are typical of early peak accents. In other words, accented syllables produced with early peaks in German are perceived as low not high. Based on this Rathcke and Harrington decided to follow the phonological analysis as proposed in GToBI, i.e. HL*.

The current study involves a somewhat similar experiment to the one performed by Rathcke and Harrington (2007). Different from the stimuli in Rathcke and Harrington (2007)'s experiment which only consisted of names with initial stress, the stimuli in the current study consist of names with various metrical structures. This allows us to better capture the full accent and how it behaves under different metrical conditions: Especially for the early peak accent which involves tonal targets preceding the accented syllable, it would be interesting to observe metrical conditions that allow for pre-accentual syllables like penultimate and final stressed words. Rathcke and Harrington (2007)'s experimental setup with only initial stressed names, does not allow us to tease apart leading tones from initial boundary tones, for instance.

Like Rathcke and Harrington (2007) we examine the scaling and alignment of two different falling patterns in German as well as Dutch. Interestingly, in contrast to Caspers (2000), Rathcke and Harrington's representation of early and late peak accents suggests a difference in peak alignment rather than peak height.

4.3 Reduction Effects

Besides studying the phonetic and phonological forms of the early and late peak in German and Dutch, this study also investigates reduction effects in early peak accents. Early peak accents presumably involve pre-accentual tonal targets (i.e. an H leading tone or boundary tone). The late peak accent on the other hand, arguably does not involve pre-accentual tonal targets and therefore the expectation would be that the two accents should

behave differently in those phonological environments lacking pre-accentual syllables. Whenever an intonational unit cannot be fully realized because of a decrease in duration or voiced segments, its realization can be subject to categorical truncation or gradient compression. Most research on tonal truncation and compression has focused on the final part of intonational accents (Rietveld and Gussenhoven 1995; Ohl and Pfitzinger 2009). This study explores reduction effects in the initial part of the early peak accents in Dutch and German, which features an f0 peak before or early in the accented syllable. This is in contrast to late peaks in which the f0 peak is observed in the middle of the accented syllable. Since the early peak appears to involve tonal targets in pre-accentual syllables, we investigate what happens when pre-accentual syllables are not available. To my knowledge no research has explored reduction effects in the initial part of this specific nuclear accent. Grice (1995) argues that leading tones undergo phonetic reduction in phrase-initial position. However, as of yet no empirical study has been done to confirm this claim.

The question we would like to answer is: Are early peaks reduced in contexts lacking pre-accentual syllables? And if so how are they reduced? The reason for investigating reduction effects is because we do not know how the falling accent would behave in reduction contexts, and since this is an exploratory research we would like to know in order to gain a more comprehensive understanding of the early vs. late peak contrast.

Reduction contexts can also tell us something about possible phonological representations. If the early peak accent is represented as "%H !H*+L" (as Caspers does), reducing the number of pre-accentual syllables will likely only compress the beginning of the accent. If on the other hand one represents the early peak accent as "H+!H*" (as Gussenhoven et al. (2005) does), then there is substantial reason to believe that the distinction between early and late peaks may be (partially) neutralized: Leading tones could be truncated in contexts lacking pre-accentual syllables. If the high leading tone in H+!H* were to be truncated one would get an non-downstepped H tonal target since there is no preceding high tone inducing downstep.

4.4 Research Questions

In this study we investigate the phonological specification and phonetic realization of early and late peak accents by means of an imitation task where the two accents are produced under different metrical conditions. As we have seen in the previous sections, although the analysis of the late peak accent is more or less similar across different account (i.e. a high tonal target followed by either a low trailing or boundary tone, H*+L/H* L%), various representations have been proposed for the early peak accent. Some of these representations suggest a difference in peak alignment between the two accents, e.g. Rathcke and Harrington (2006)'s H+L* representation for the early peak accent, while others suggest a difference in peak height, e.g. Caspers (2000)'s !H*+L representation.

There are several research questions that will be addressed in this chapter. First of all, what would be the best phonological representation to capture the two falling accents, especially the early peak accent? This general research question can be divided into three sub-questions. We will first focus on the tonal targets that are associated with accented syllables: Should the early peak accent be analyzed as a downstepped H target (like Caspers (1999)'s !H*L and Gussenhoven et al. (2005)'s H+!H* analysis), or should it be represented as an L target(like Rathcke and Harrington (2006)'s H+L* analysis for German)? If the early peak accent contains a downstepped H starred tone the expectation is that there is a peak preceding this downstepped H target and the lowest point in pitch has not been reached yet. If on the other hand the early peak accent consists of a L starred tone then the expectation is the lowest point in pitch is reached in the accented syllable and there is no lowering of pitch in syllables following the accented syllable. For the late peak the consensus is that it is realized with a high tonal target in the accented syllable.

Secondly, the accent onset will be examined: Are there high leading tones (e.g. Gussenhoven et al. (2005) for Dutch and Rathcke and Harrington (2006) for German) or high initial boundary tones (e.g. Caspers (2000)) in the early peak accent. With high leading tones a peak is expected in the syllable preceding the accented syllable whereas with high initial

boundary tones a gradual decline in pitch is expected from the onset of the word till the starred tone.

Finally, we will research whether there is empirical evidence for trailing tones in the early and late peak accents: Caspers assumes low trailing tones for both accents, whereas Grice et al. (2006) and Rathcke and Harrington (2006) assume a low boundary tone for late peak accents. If there are low trailing tones the lowest point in pitch is expected in the syllable following the accented syllable. If on the other hand there is only a low boundary tone, a gradual decline from the H target tone to the end of the word is expected.

A second research question that will be discussed in this chapter, has to do with reduction effects: Do the early and late peak accents show the same reduction effects? Given that in most analyses the early peak accent is represented with leading tones, whereas the late peak accent is not, the expectation would be that the two accents will show different reduction effects in contexts lacking pre-accentual syllables: The early peak accent will arguably be more affected in such contexts than the late peak accent.

To answer these research questions we will be looking into the phonetic realizations of these different accents under various metrical contexts (e.g. monosyllabic vs. trisyllabic and initial vs. penultimate vs. final stress patterns) in a systematic way. This allows us to gain a more comprehensive understanding of the early vs. late peak contrast.

4.5 Experiment 1.1: Dutch

An online production experiment was set up in the form of an imitation task. Both Dutch (Experiment 1.1) and German were tested (Experiment 1.2) following a similar experimental set up.

4.5.1 Participants

A total of 20 participants were recruited for the experiment using the online platform Prolific (www.prolific.co). All of them were native speakers of Standard Dutch as spoken in the Netherlands. Participants received a compensation of £2.5 for a 20 minute experiment.[5] All of them were naive as to the purpose of the experiment.

4.5.2 Auditory Stimuli

The example dialogue to be imitated consisted of a question and an answer. There were only two example dialogues: one for the early peak accent (47), and one for the late peak accent (48). The dialogues were constructed based on the idea that early peak accents are consistent with something that is expected, whereas late peak accents are more neutral and there is no such expectation (Caspers 2000). The answer consisted of a disyllabic proper name with final stress. Both example dialogues have been recorded by the author. Each answer was produced 25 times, i.e. 25 productions of 'Heleen' with an early peak and 25 with a late peak.

(47) **Example dialogue early peak:**

Iedereen was op tijd, maar wie was weer te laat? He<u>leen</u>
Everyone was on time, but who was again too late Heleen

(48) **Example dialogue late peak:**

Wie was te laat? He<u>leen</u>
who was too late Heleen

To build prototypical representative stimuli, the mean pitch at different time points of all 25 productions has been used to create a sound with a synthesized pitch trajectory: 20 evenly distanced time points for the first syllable 'he' and 60 for the second syllable 'leen' (see Figures 4.6a and 4.6b). It was noted that the intensity peak in the accented syllable was higher than the preceding syllable for the late peak, but the opposite pattern was true

[5]This comes down to £7.5/h.

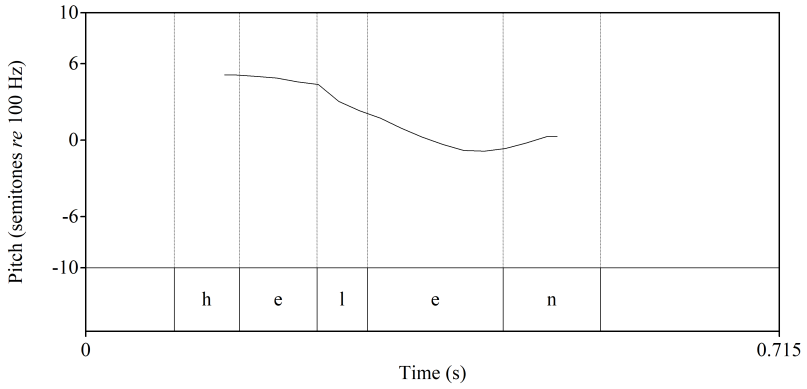
for the early peak accent. To make sure there was no intensity difference between the accented and unaccented syllable, the intensity peaks in both syllables have been made equal by scaling the intensity peak in the first syllable to match the intensity peak in the second syllable. In addition, the vowel durations have been manipulated to match the mean vowel durations of the 25 productions.

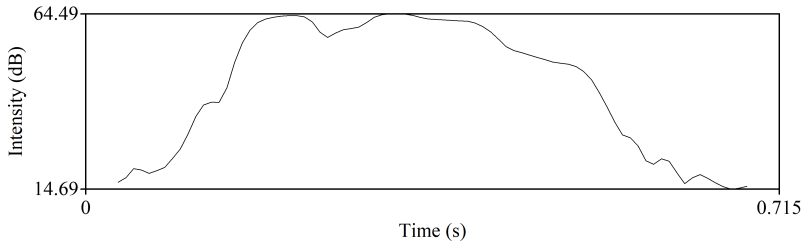
4.5.3 Target Stimuli

The target stimuli to be recorded by the participants were similar to the example dialogue. The answers consisted of either a monosyllabic name or a trisyllabic name differing in stress pattern (see Example 49).

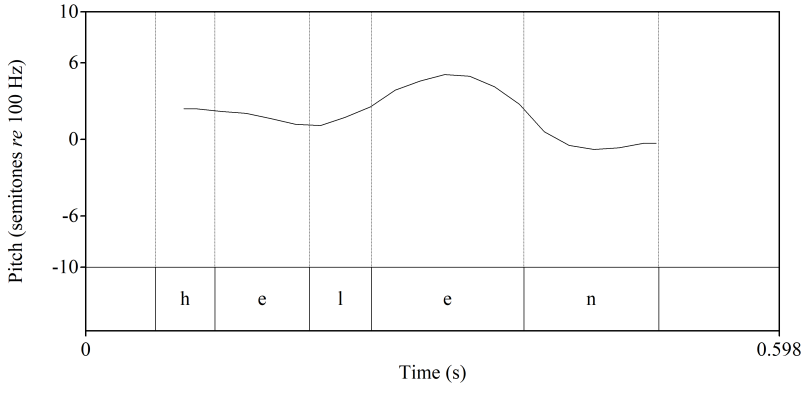
Since the early peak accent is said to have its f0-peak early in the accented syllable or in the preceding syllable, we are interested in knowing what would happen when no such preceding syllable was present. Therefore, stimuli with initial stress and monosyllabic names were created (Examples 49a, 49b). In addition, names with initial stress should allow us to distinguish trailing tones from final boundary tones: With trailing tones the lowest point in pitch is expected in the syllable following the accented syllable, whereas with low boundary tones a gradual decline starting from the peak to the end of the word is expected. The stimuli with penultimate stress are chosen since they would provide enough metrical material for the accents to be realized without much reduction, i.e. the accented syllable with a potential leading tone in the preceding syllable or a trailing tone in the following syllable. We also looked at names with final stress as we are interested in determining whether we are dealing with initial boundary tones or leading tones in the early peak accent: If there are high leading tones then a peak is expected in the syllable preceding the accented syllable. If there are high boundary tones, a gradual decline starting from the onset of the word till the accented syllable is expected.

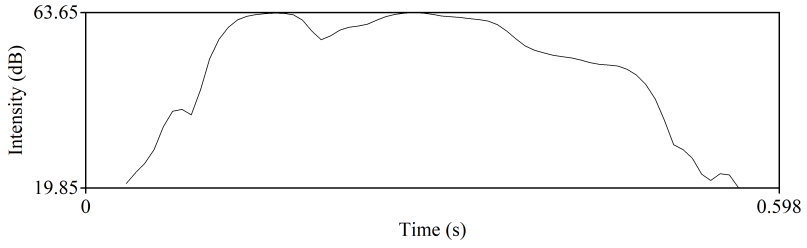
(49) Iedereen was nuchter, maar wie was weer dronken? /Wie was dronken?
Everyone was sober, but who was again drunk /who was drunk





(a) Early peak accent





(b) Late peak accent

Fig. 4.6: Pitch trajectory of the example answers to be imitated. The pitch shown here is the result of taking the mean pitch at different points of all the 25 productions.

a. Lien (monosyllabic)

b. Lieneke (trisyllabic,initial stress)

c. Joliene (trisyllabic, penultimate stress)

d. Marjolien (trisyllabic, final stress)

A total of 16 item sets were created and each condition is repeated four times. Given that there are two accent conditions and four metrical conditions, each participant would produce a total of 32 accents (i.e. 2 accents*4 conditions * 4 repetitions = 32)

4.5.4 Procedure

Participants were redirected to the online experiment via Prolific. They were presented with a small dialogue and asked to imitate a similar dialogue as closely as possible in intonation. The example question and answer to be imitated were presented both in text form as well as auditorily. In order to make the participant more familiar with the task and the intonation, a training phase preceded the main blocks.

The experiment was made using jsPsych (De Leeuw 2015) and the scripts were created by the Prosodylab at McGill University. Four blocks were created, two of which were part of the actual experiment and two other blocks were training blocks. One main block consisted of only early peak accents, and the other block consisted of only late peak accents. Each main block was preceded by a training phase consisting of 4 trials. The main experiments contained 32 trials each, thus making the total number of trials a participant had to loop through, 72 trials. Eight experimental lists were constructed by rotating through the 8 conditions, i.e. the number of accents in the answer (single accent or two accents) and the stress pattern of the names (initial, penultimate, final or NA for monosyllabic names). Each condition was repeated four times, thus there were four items per condition. Half of the participants were presented with the early peak accents first and the other half with the late peak accents first. Participants were randomly assigned to one of these lists.

4.5.5 Statistical Analysis

For the results and statistical analyses the pitch trajectories produced by the participants were observed (see Figure 4.8 for the mean pitch trajectories across the different conditions). In order to be able to perform any statistical analyses on these pitch trajectories, the f0 accents were first modeled as continuous trajectories by means of Functional Data Analysis (FDA) (Ramsay et al. 2009; Gubian 2013) and subsequently a principle component analysis (PCA) was performed. Functional data analysis allows for an analysis of information on curves or functions. It transforms a certain curve into a spline[6] function, thus making it possible to perform data analysis on its parameters. The reason for using FDA is to avoid having to deal with finding peaks and elbows in smooth trajectories or lacking discrete turning points. The central idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while at the same time retaining as much as possible of the variation present in the data set. It does this by transforming the data into fewer dimensions, which act as summaries of features. In this section, a more in-depth PC analysis is only provided for the monosyllabic cases. The PCA figures for other conditions can be found in the appendix. For the FDA late peak accents were compared to early peak accents.

For each metrical condition a total of 160 recordings were obtained (2 accents * 4 repetitions * 20 participants = 160), a total of 640 recordings were obtained (2 accents * 4 repetitions * 4 metrical conditions * 20 participants = 640). One participant had to be excluded since no recordings were made and other recordings had to be excluded due to poor audio quality, the recording being cut-off or the participant producing a wrong target, leaving us with a total of 608 recordings. Out of the 608 recordings 17 were produced with two pitch accents within the same word (see Figure 4.7). I will discuss these specific examples in the final discussion session. A total of 591 recordings were used for data analysis.

[6]Splines are complex functions that add curves together to make a continuous and irregular curves

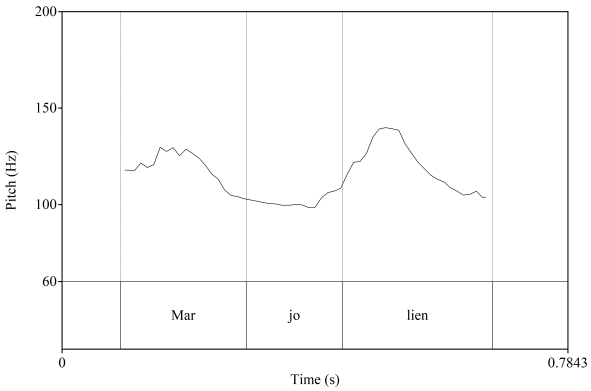


Fig. 4.7: Trisyllabic name with two pitch accents within the same word. An A accent on *mar* and *lien*.

To see whether the observed differences are statistically significant several linear mixed-effects regression model (LMM) were run using R (Team 2008) and the LMER function in the lme4 Rpackage (Bates et al. 2014). The principle component scores were taken as outcome variable and as fixed effects *accent* was added to the model. As random effects, by-subject random intercepts were included in the model.

4.5.6 Results

Figure 4.9 shows the accents captured by the different principle components, these figures only show the accents for *Lien*, for the other words see Appendix B. For reasons of space, only the figures for *Lien* have been provided in this chapter, all the other figures can be found in the appendices at the end of this chapter. As can be seen from Figure 4.9a, Principle Component 1 for the monosyllabic *Lien* best describes the slope of the accent, and accounts for 53.2% of the variance. Principle Component 2 seems to capture the flatness of the accent, and accounts for 35.1% of the variance (see Figure4.9b). Less clear is what dimension is captured by Principle Component 3, but the accent that it captures is shown in Figure 4.9c. From Figure 4.10 it becomes clear that the two accents are best categorized by PC2.

It should be remarked that the principle components do not capture the exact same phenomenon across different words, the dimensions they capture heavily depend on the

input data and will therefore differ per word. There does however, seem to be a general phenomenon that the different principle components capture across the different metrical conditions: Broadly speaking, both PC1 and PC2 seem to capture the peak alignment and peak height, less clear is which phenomenon is captured by PC3.

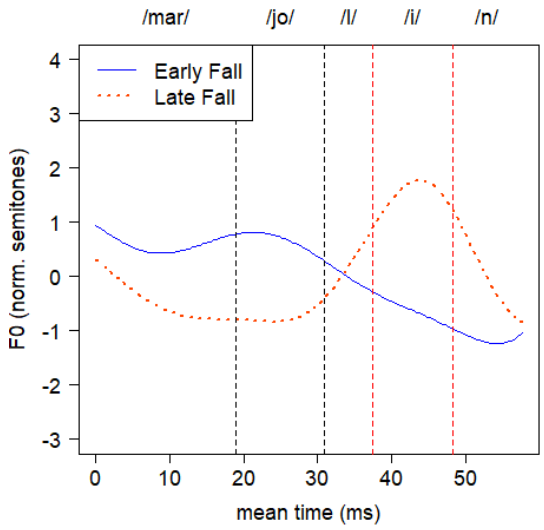
For the monosyllabic *Lien*, a model with *accent* (AIC 813.64) did significantly improve over an intercept-only null model (AIC 815.61) for PC1 (χ^2 (1, N = 149) = 3.9673 , p = 0.04639). Although *accent* improved the model, the model itself reported no significantly difference between the early peak accent and the late peak accent based on their PC1, i.e. the slope of the accent. For PC2 there was a highly significant effect of adding *accent* to the model (AIC 750.78) over an intercept-only null model(AIC 777.25) (χ^2 (1, N = 149) = 28.476 , p = 9.485e-08): The early peak accent is significantly flatter than the late peak accent (beta = 6.76, 95% CI [4.98, 8.55], t(143) = 7.49, p < .001).

For the initial stress condition *Lieneke*, a model with *accent* (AIC 861.73) did significantly improve over an intercept-only null model (AIC 867.10) for PC1 (χ^2 (1, N = 149) = 4.4158 , p = 0.03561): Although *accent* improved the model, the model itself reported no significantly difference between the early peak accent and the late peak accent based on their PC1. For PC2 there was a highly significant effect of adding *accent* to the model (AIC 891.72) over an intercept-only null model(AIC 911.36) (χ^2 (1, N = 149) = 21.643 , p = 3.284e-06): The peak in the early peak accent is lower and more to the left than the peak in the late peak accent (beta = 6.02, 95% CI [3.94, 8.10], t(143) = 5.72, p < .001). Finally, for PC3 adding *accent* to the model (AIC 861.73) significantly improved the null model (AIC 867.10) (χ^2 (1, N = 149) = 7.3741 , p = 0.006617): The early peak accent is significantly different from the late peak accent in terms of PC3 (beta = -2.57, 95% CI [-4.64, -0.50], t(143) = -2.45), it is not exactly clear what is captured by PC3.

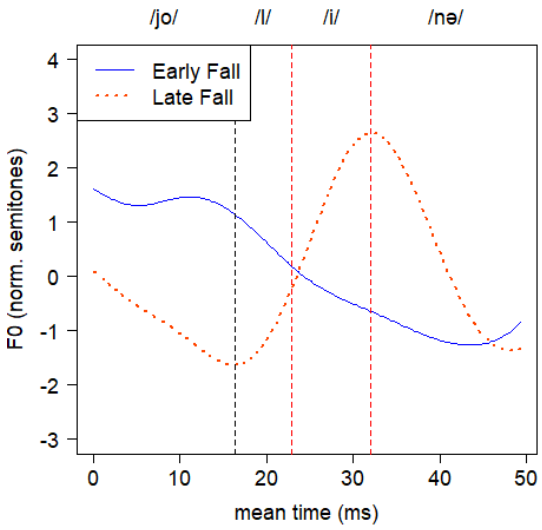
For the penultimate stress condition *Joliene*, a model with *accent* (AIC 899.73) did significantly improve over an intercept-only null model (AIC 923.25) for PC1 (χ^2 (1, N = 142) = 25.512 , p = 4.396e-07): The early peak accent is significantly different from the late peak

accent based on their PC1: The peak in the early peak accent is aligned earlier and is lower than in the late peak accent (beta = -14.56, 95% CI [-19.07, -10.06], t(136) = -6.40, p < .001). For PC2 there was a significant effect of adding *accent* to the model (AIC 826.27) over an intercept-only null model(AIC 828.43) (χ^2 (1, N = 142) = 4.162 , p = 0.04134), but the model self did not report a significant result. Finally, for PC3 adding *accent* to the model (AIC 726.05) significantly improve the null model (AIC 733.48) (χ^2 (1, N = 142) = 9.4269 , p = 0.002138): The early peak accent is significantly different from the late peak accent in terms of PC3 (beta = 2.62, 95% CI [0.93, 4.30], t(136) = 3.07, p = 0.003).

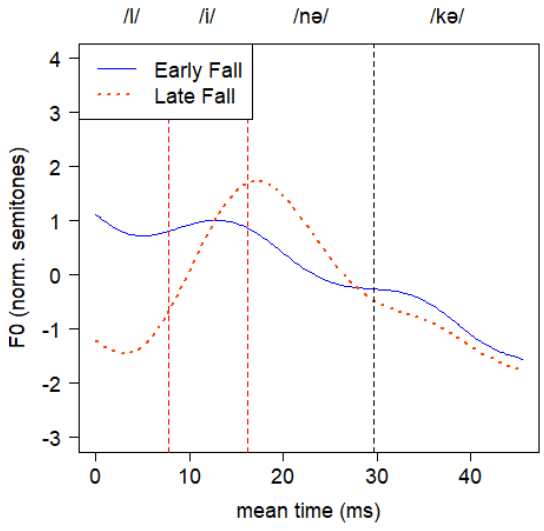
For the final stress condition *Marjolien*, a model with *accent* (AIC 969.47) did significantly improve over an intercept-only null model (AIC 987.35) for PC1 (χ^2 (1, N = 151) = 19.883 , p = 8.233e-06): The peak in the early peak accent is aligned earlier and is lower than in the late peak accent (beta = 10.82, 95% CI [6.54, 15.10], t(145) = 4.99, p < .001). No other significant effects were found.



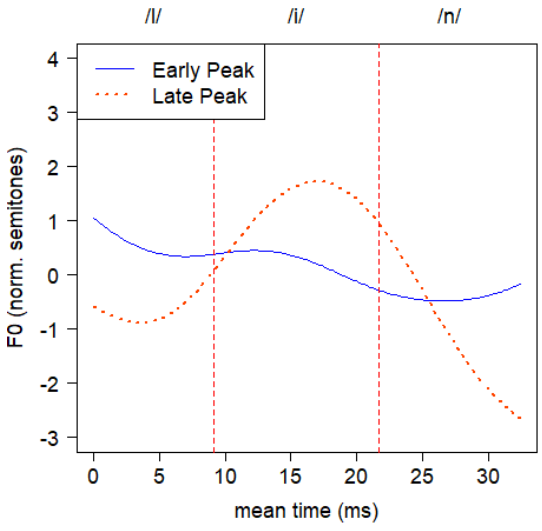
(a) Final stress



(b) Penultimate stress

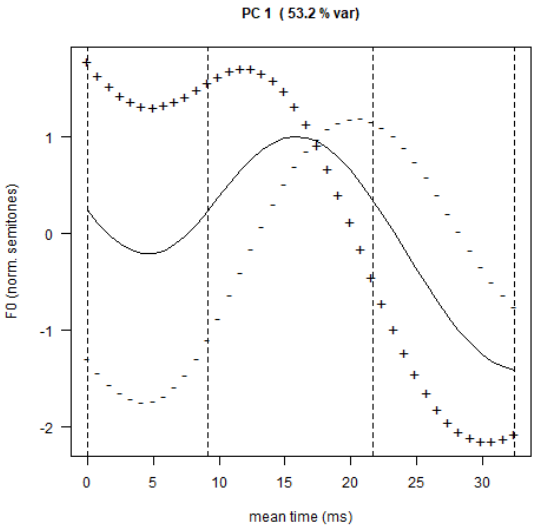


(c) Initial stress

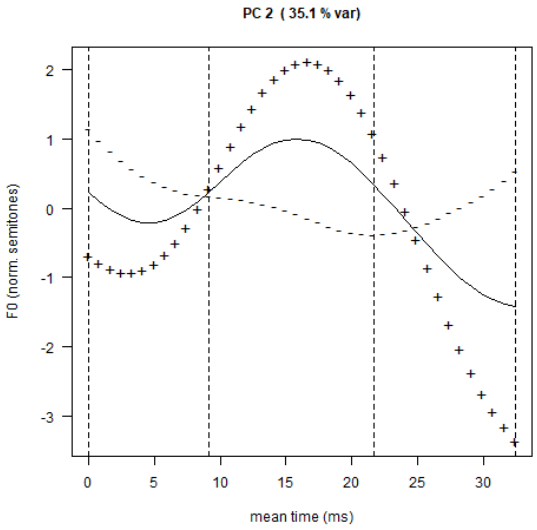


(d) Monosyllabic

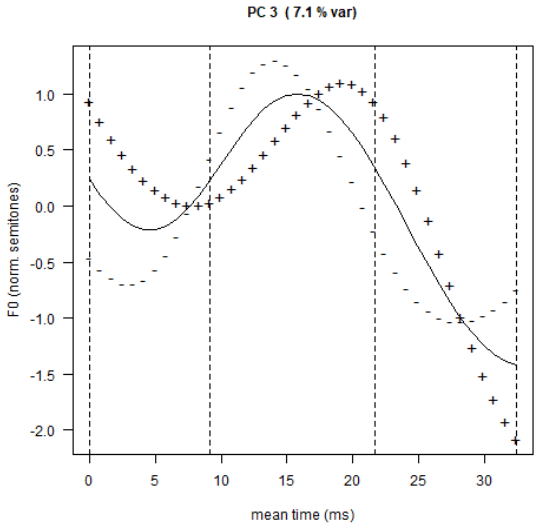
Fig. 4.8: Mean normalized f0-curves for both accents. The vertical black line indicates a syllable boundary, the vertical red line indicates the boundaries of the vowel in the accented syllable.



(a) Principle Component 1 *Lien*



(b) Principle Component 2 *Lien*



(c) Principle Component 3 *Lien*

Fig. 4.9: The solid line displays the mean accent captured by a certain principle component over the course of the mean duration of a certain word. The line consisting of plus symbols represents the pitch trajectory captured by higher values of a certain principle component and the minus line lower values.

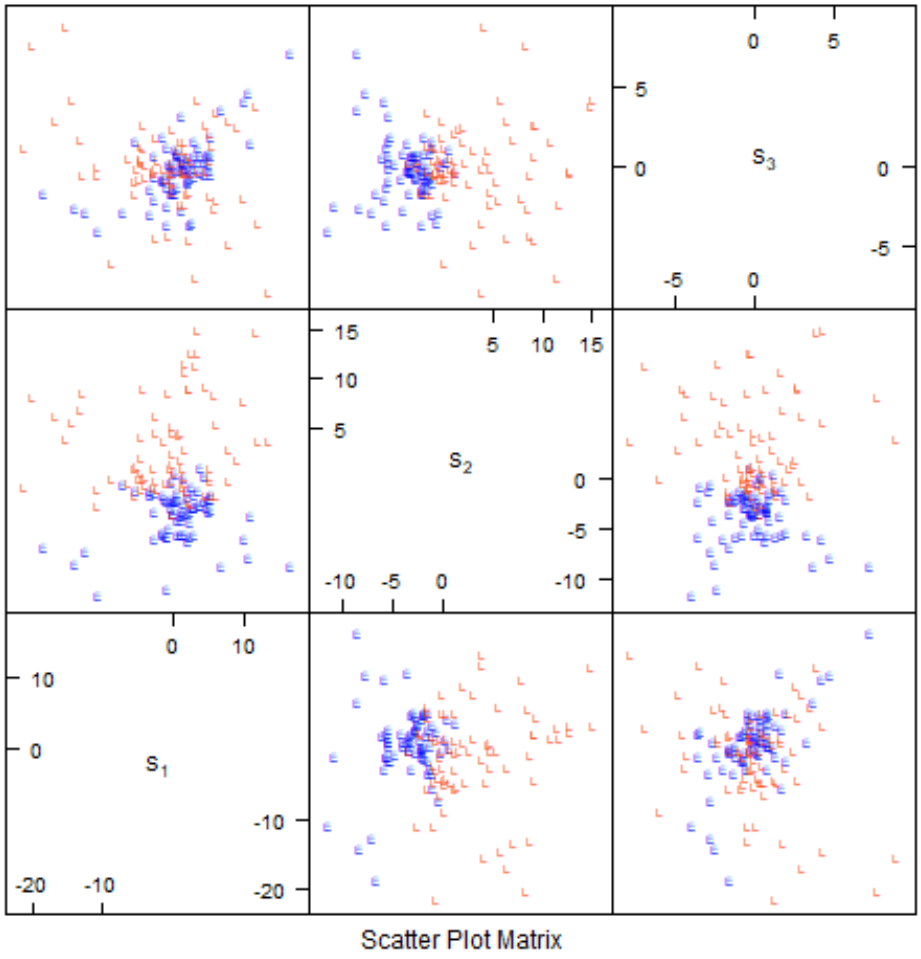


Fig. 4.10: PCA 1, 2 and 3 plotted against each other for *Lien*. Late peaks are represented by red *L*s and early peaks by blue *E*s. The labels S_1, S_2, and S_3 correspond to PC1, PC2 and PC3, respectively.

4.5.7 Conclusion

Based on the results from experiment 1.1 it can be concluded that in Dutch there are clear phonetic differences between the early peak and late peak accents across different metrical conditions. The various metrical conditions do not all differ along the same PCA dimensions as becomes clear from the fact that different principle components capture different dimensions and not the same principle components are significant across the different metrical conditions.

What becomes clear from the principle component analyses is that the monosyllabic name *Lien* seems to be different from the other metrical conditions in terms of along which dimension the two accents differ: The early peak accent in monosyllabic names seem to be flatter in pitch than the late peak accent in monosyllabic names. For the other metrical conditions the difference between the two accents seems to be that the peak in the early peak accent is lower and timed earlier than the peak in the late peak accent. To get a better idea of along which dimensions the two accents differ for a specific metrical condition in more detail, see Appendix B.

In the final section, the specific pitch trajectories and what phonological representation goes best with these accents will be discussed in more detail: It will be argued that the early peak accent is best described as being part of a accent "%H H+!H* L%" (i.e. a high boundary tone followed by a high leading tone, a downstepped high starred tone and a final low boundary tones), and the accent containing the late peak accent as "H* L%" (i.e. a high starred tone followed by a low boundary tone). We will also discuss why tonal targets in the early peak accent are compressed (i.e. results in a relatively flat trajectory), whereas in the late peak tonal targets seem to be more distinctly realized.

4.6 Experiment 1.2: German

Experiment 1.2 follows the same procedure as Experiment 1.1 and is therefore not repeated in this section.

4.6.1 Participants

A total of 20 participants were recruited for the experiment using the online platform Prolific (www.prolific.co). All of them were native speakers of Standard German as spoken in

Germany. Participants received a compensation of £2.5 for a 20 minute experiment.[7] All of them were naive as to the purpose of the experiment.

4.6.2 Auditory Stimuli

The example dialogue to be imitated consists of a question and an answer: There are only two example dialogues one for the early peak (50) and one for the late peak (51). The answer consists of a disyllabic proper name with final stress. Both example dialogues have been recorded by a native speaker of German. Each answer was produced 20 times, i.e. 20 productions of 'Marie' with an early peak and 20 with a late peak.

(50) **Example dialogue early peak:**

Alle waren pünktlich, aber wer war wieder zu spät? Marie
Everyone was on.time, but who was again too late Marie

(51) **Example dialogue late peak:**

wer war zu spät? Marie
who was too late Marie

4.6.3 Target Stimuli

The target stimuli to be recorded by the participants were similar to the ones in Dutch (see Example 52), except that no trisyllabic with initial stress could be found, instead a bisyllabic word with initial stress was used.

(52) Alle waren nüchtern, aber wer war wieder betrunken?
Everyone was sober, but who was again drunk

a. Lin (monosyllabic)

b. Lina (disyllabic,initial stress)

c. Alina (trisyllabic, penultimate stress)

d. Evelin (trisyllabic, final stress)

[7] This comes down to £7.5/h.

4.6.4 Statistical Analyses

Figure 4.11 shows the mean pitch trajectories across the different conditions in German. To see whether there were any significant differences, several linear mixed-effects regression model (LMM) were run using R (Team 2008) and the LMER function in the lme4 R package (Bates et al. 2014). The outcome variable was the PC score and as fixed effect *accent* was added to the model. As random effects, by-subject random intercepts were included in the model. A total of 640 recordings were obtained (2 accents * 4 repetitions * 4 metrical conditions * 20 participants = 640). Some recordings had to be excluded due to poor audio quality, the recording being cut-off or the participant producing a wrong target (e.g. producing a name with initial stress instead of with final stress), leaving us with a total of 532 recordings. Out of the 532 recordings 3 were produced with two pitch accents within the same word and were therefore excluded. A total of 529 recordings were used for data analysis.

4.6.5 Results

For the monosyllabic *Lin*, a model with *accent* (AIC 1011.4) did not significantly improve over an intercept-only null model (AIC 1012.9) for PC1 (χ^2 (1, N = 151) = 3.4913 , p = 0.06169). For PC2 there was a highly significant effect of adding *accent* to the model (AIC 869.64) over an intercept-only null model(AIC 891.87) (χ^2 (1, N = 151) = 24.229 , p = 8.552e-07): The peak in the early peak accent is flatter than the peak in late peak accent (beta = 5.48, 95% CI [3.86, 7.10], t(145) = 6.69, p < .001).

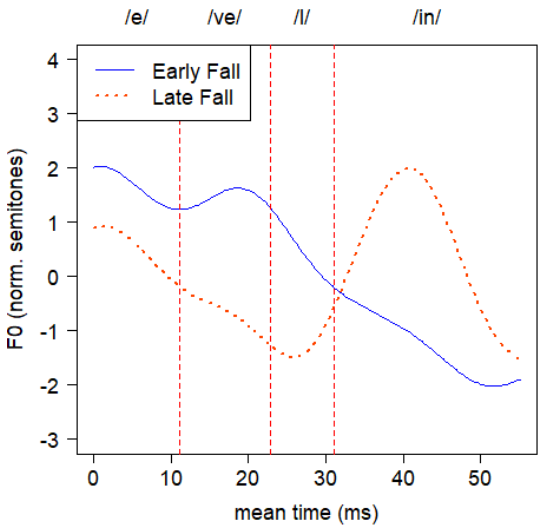
For the initial stress condition *Lina*, the model with *accent* (AIC 1039.5) did significantly improve over an intercept-only null model (AIC 1041.8) for PC1 (χ^2 (1, N = 146) = 4.2354 , p = 0.03959), but the model itself reported to significant result. For PC2 there was a highly significant effect of adding *accent* to the model (AIC 905.44) over an intercept-only null model(AIC 930.28) (χ^2 (1, N = 146) = 26.835 , p = 2.216e-076): The peak in the early peak accent is flatter than the peak in late peak accent (beta = 6.92, 95% CI [5.00, 8.83], t(140)

= 7.13, p < .001;). Finally, for PC3 adding *accent* to the model (AIC 768.10) significantly improves the null model as well (AIC 780.04) (χ^2 (1, N = 149) = 25.764 , p = 3.858e-07): For the late peak accent the peaks are higher and the valleys lower than the early peak accent (beta = -2.74, 95% CI [-3.77, -1.71], t(142) = -5.26, p < .001).

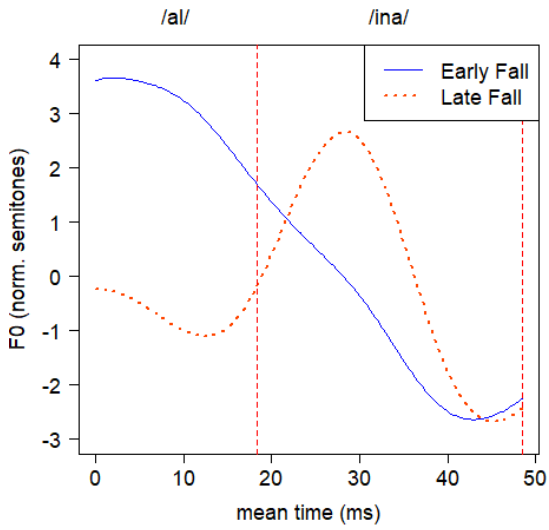
For the penultimate stress condition *Alina*, a model with *accent* (AIC 1079.2) significantly improves over an intercept-only null model (AIC 1103.7) for PC1 (χ^2 (1, N = 149) = 26.514 , p = 2.616e-07): The peak in the early peak accent is aligned earlier and is higher than in the late peak accent (beta = -18.15, 95% CI [-23.76, -12.54], t(143) = -6.40, p < .001). Similarly, for PC2 there was a significant effect of adding *accent* to the model (AIC 1002.3) over an intercept-only null model(AIC 1009.9) (χ^2 (1, N = 149) = 9.5756, p = 0.001972): The overall accent is flatter in the early peak accent than in the late peak accent (beta = 4.05, 95% CI [1.27, 6.83], t(143) = 2.88, p = 0.005).

For the final stress condition *Evelin*, a model with *accent* (AIC 589.81) did significantly improve over an intercept-only null model (AIC 600.68) for PC1 (χ^2 (1, N = 83) = 12.866 , p = 0.0003347)[8]: The peak in the early peak accent is aligned earlier and is lower than in the late peak accent (beta = -11.47, 95% CI [-17.89, -5.05], t(77) = -3.56, p < .001). For PC2 there was a small significant effect of adding *accent* to the model (AIC 529.50) over an intercept-only null model (AIC 531.57) (χ^2 (1, N = 83) = 4.0737 , p = 0.04356), but the model itself reported no significant results. Finally, for PC3 adding *accent* to the model (AIC 469.07) also significantly improve the null model (AIC 477.32) (χ^2 (1, N = 83) = 10.254 , p = 0.001364): The early peak accent is significantly different from the late peak accent in terms of PC3 (beta = 3.76, 95% CI [1.72, 5.81], t(77) = 3.67, p < .001).

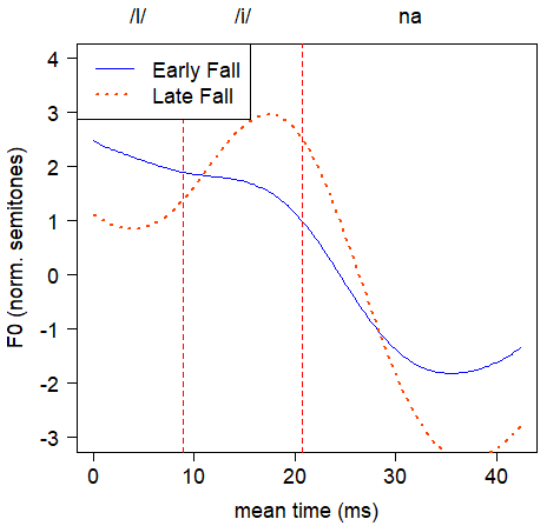
[8] Almost half of the recordings had to be removed since they were produced with an initial stress rather than a final stress.



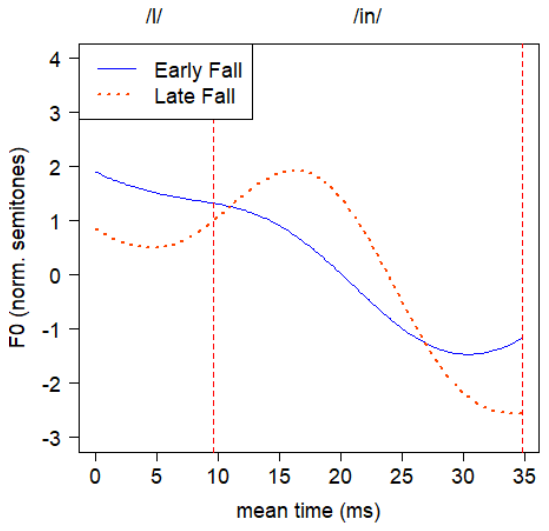
(a) Final stress



(b) Penultimate stress



(c) Initial stress



(d) Monosyllabic

Fig. 4.11: Mean normalized f0-curves for both accents. The vertical black line indicates a syllable boundary, the vertical red line indicates the end of the vowel in the accented syllable.

4.6.6 Conclusion

Similar to Experiment 1.1, Experiment 1.2 shows that there are clear phonetic differences between the early peak and late peak accents in German across different metrical conditions. Again, the various metrical conditions do not all differ along the same PCA dimensions as became clear from the fact that different principle components capture different

dimensions and not the same principle components are significant across the different metrical conditions.

What becomes clear from the principle component analyses is that monosyllabic names and names with initial stress behave similarly: The peaks in the early peak accents are flatter than in the late peak accent. For the other metrical conditions the difference between the two accents seems to be that the peak in the early peak accent is timed earlier than the peak in the late peak accent. To get a better idea of along which dimensions the two accents differ for a specific metrical condition in more detail, see Appendix B.

4.7 Discussion

Based on the results it can be concluded that the two falling accents are phonetically different: For all of the metrical conditions, the two accents significantly differed at least in one principle component in both German and Dutch. The two languages behave surprisingly similar, the only characteristic that is slightly different is that there is more of a declining trend over the course of the word in German than in Dutch as can be seen from the mean f0 trajectories in the two languages. More specifically, the PCA results tell us that for the penultimate and final stressed names, the peaks in the early peak accent are timed earlier in the word and tend to be lower than the peaks in the late peak accent.[9] For the monosyllabic and initial stressed names, the PCA results generally indicate that peaks were flatter (or almost non-existent) in the early peak accent than in the late peak accent.

To answer our research question concerning how the accents are best represented, we will have to observe the mean pitch trajectories (Figures 4.11 and 4.8) as the PCA results tell us little about the "correct" representation. With regards to the late peak accent and the question whether there is a low trailing tone or a low final boundary tone, we can conclude that there is a low final boundary tone rather than a low trailing tone (i.e. H* L%):

[9]It should be noted that only for the penultimate name in German was the peak in the early peak accent actually higher than in the late peak accent. I do not have a specific explanation for this result.

There is interpolation from the f0-peak to the end of the word, rather than a low target in the syllable following the accented syllable (see Figures 4.8c and 4.11c).[10] This is in agreement with Rathcke and Harrington (2006) and Grice et al. (2006) and in disagreement with Caspers (2000) who represented the late peak as H*+L.

Let us now turn to the issue of how to best represent the early peak accent phonologically. In terms of the tonal target, a downstepped H tone would best capture the results: This becomes especially clear from Figures 4.8b and 4.11b where the lowest f0-peak is not reached until the next syllable. In terms of the accent onset, the final stressed names provide evidence for a high leading tone, the f0-peak occurs in the syllable preceding the accented syllable (see Figures 4.8a and 4.11a). The results do not support an analysis with a high initial boundary tone instead of a leading tone, if this were the case one would have expected a gradual downward trend in pitch from the beginning of the word all the way to the end of the word, because of interpolation of pitch between an initial boundary tone and a the target tone. Instead the high pitch elbow on the syllable preceding the accented syllable is best explained by a high leading tone.

However, there is also evidence for a high initial boundary as becomes apparent from the high onset in the early peaks for all trisyllabic names, so one can argue that there is both a high initial boundary tone as well as a high leading tone. The early peak then is best transcribed as %H H!H* L%, i.e. a high initial boundary tone followed by a high leading tone preceding the accented syllable, a downstepped high tone on the accented syllable and a low final boundary tone. Alternatively one could argue that the high boundary tone spreads up to the accented syllable, this would explain the high initial plateau and there would be no need to posit a bitonal accent. Still this would not explain why there is a clear peak in the syllable preceding the accented syllable (see Figures 4.8a and 4.11a). Therefore an analysis with both a high initial boundary tone and high leading tone seems to best capture the accent. Representing the early peak accent as H+!H* is in agreement

[10] It should be noted that for German no trisyllabic name with initial stress could be found, so there is no direct evidence for a low final boundary rather than a trailing tone in German.

with Gussenhoven et al. (2005)'s representation of this accent. Rathcke and Harrington (2006; 2007; 2010)'s arbitrary choice of representing the early peak accent as H+L* is not supported the mean pitch plots.

One important thing to note is that the proposed representations for the two falling accents (i.e. H+!H* for the early peak accent and H* for the late peak accent) cannot explain why in the case of the early peak accent, a H (leading) tone would be lower than a H (starred) tone in the late peak accent. It might be that H tones in accented syllables are simply produced more distinctly. Nonetheless, I have no satisfying explanation for why this would be the case and I leave it up for future research it provide an adequate answer.

Having established that the early peak accent contains a high leading tone which tend to occur on preaccentual syllables, one can now wonder what happens when there are no preaccentual syllables. Comparing the f0-curve of trisyllabic names with penultimate and initial stress (Figures 4.8b and 4.8c in Dutch, and 4.11b and 4.11c in German), it becomes clear there is little to no reduction of the initial part of the nuclear accent in early peaks. In other words, the lack of pre-accentual syllables does not lead to the reduction of tones, rather the overall accent is placed on top of the available segments. In German, the monosyllabic names show a similar pattern to the trisyllabic initially stressed names in the sense that there is little to no reduction and the tonal gesture seems to be temporally compressed rather than heavily reduced in shape (see Figure 4.11d).

In Dutch however, there is a noticeable difference between trisyllabic names with initial stress and monosyllabic names (4.8d). Although we can conclude that early peaks are not severely reduced in contexts lacking pre-accentual syllables, one does observe clear reduction effects in monosyllabic words hosting three tonal targets (H, !H*, and L%), where the early peak exhibits a relatively flat pitch trajectory, with a lower onset relative to German. One could argue that this is due to phonetic undershoot of both the H leading tone and L final boundary tone. However one wonders why in late peaks, speakers are able to produce a rather more complex (rising-falling) pitch trajectory. In contrast to early peaks,

late peaks are more distinctly realized in monosyllabic words. This shows that different nuclear configurations of similar tonal complexity are subject to different degrees of reduction under time pressure and in different ways for different languages.

One possible explanation for this discrepancy could be that perhaps vowels in early peaks are shorter than in late peaks and therefore there would not be enough time to articulate such a complex pitch trajectory. However, there is no evidence supporting the claim that vowels in early peak accents are shorter than in late peaks. In fact, a mixed-effects linear regression shows that a model with *accent* (AIC 682.57) did not significantly improve over an intercept-only null model (AIC 683.39) for vowel duration (χ^2 (1, N = 149) = 2.8181 , p = 0.09321). Therefore the conclusion can be drawn that the two accents do not significantly differ in terms of vowel duration.

Another possible explanation could come from what has been named the "Effort code" by Gussenhoven (2002; 2016b). According to the "Effort code", wider pitch range and more precise realizations of pitch events are associated with greater significance. In case of the early and late peak, only the late peak seems to be more appropriate in narrow focus, i.e. to focus one specific word. One could interpret this function of the late peak with greater significance and therefore with a wider pitch range and more precise realizations. This could potentially explain why in monosyllabic words the late peak is time compressed while the early peak is reduced in tonal complexity.

Finally, let us revert back to the question how intonation should be represented and whether there is a clear correspondence between form and function. Although not explicitly tested in this chapter, some of the tonal configurations produced by participants provide support for a more holistic approach to intonation in which a combination of tonal targets (boundary tones and pitch accents) give rise to a certain semantics rather than each tonal target itself having its own semantics. This can be clearly seen when comparing Figure 4.12 to 4.13. If there is truly a strict correspondence between meaning and pitch accent then one would not expect to find two pitch accents within the same word as is the case

in Figure 4.12: According to Jackendoff (1972) and Büring (1997) for instance each falling accent would evoke a set of alternative propositions, this analysis is problematic for a case like Figure 4.12 where two accents occur in the same word and can therefore not each invoke alternatives.

In conclusion, not only do the results from the production experiment tell us the phonetic differences between the two peak accents that are part of hat patterns, it has also provided evidence against a more compositional approach to intonational meaning. In the following chapters more semantic evidence against a compositional approach to intonational meaning will be provided.

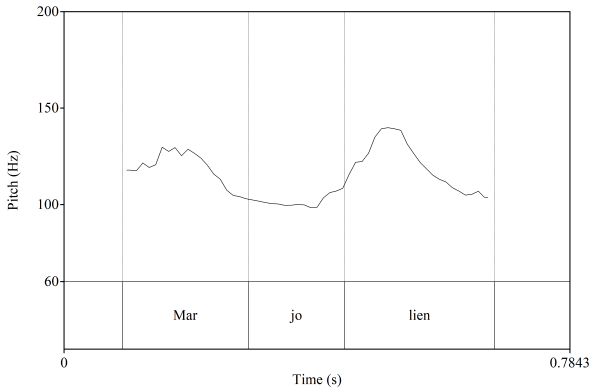


Fig. 4.12: Trisyllabic name with two pitch accents within the same word. An falling accent on *mar* and *lien*.

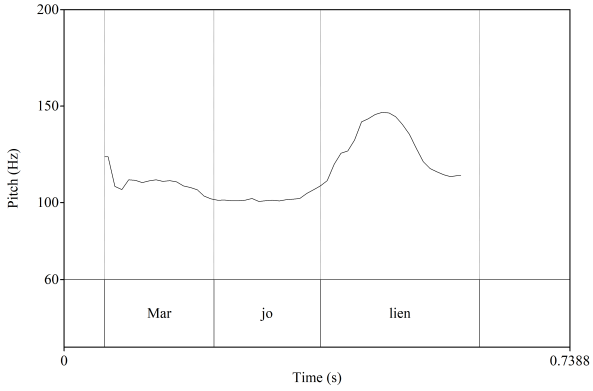


Fig. 4.13: Trisyllabic name with one pitch accents. An falling accent on *lien*.

Appendix A: Statistical Models

	PC1	PC2	PC3
(Intercept)	$0.85\ (1.11)$	$-3.41\ (0.48)^{***}$	$-0.08\ (0.35)$
accentLate Peak	$-1.75\ (1.87)$	$6.76\ (0.90)^{***}$	$0.14\ (0.57)$

$^{***}p < 0.001; {}^{**}p < 0.01; {}^{*}p < 0.05$

Table 4.1: Statistical Models Experiment 1.1: Lien

	PC1	PC2	PC3
(Intercept)	$-1.03\ (1.23)$	$-3.11\ (0.83)^{***}$	$1.30\ (0.55)^{*}$
accentLate Peak	$2.16\ (2.50)$	$6.02\ (1.05)^{***}$	$-2.57\ (1.05)^{*}$

$^{***}p < 0.001; {}^{**}p < 0.01; {}^{*}p < 0.05$

Table 4.2: Statistical Models Experiment 1.1: Lieneke

	PC1	PC2	PC3
(Intercept)	$7.43\ (1.37)^{***}$	$-0.97\ (0.78)$	$-1.27\ (0.84)$
accentLate Peak	$-14.56\ (2.28)^{***}$	$1.93\ (2.04)$	$2.62\ (0.85)^{**}$

$^{***}p < 0.001; {}^{**}p < 0.01; {}^{*}p < 0.05$

Table 4.3: Statistical Models Experiment 1.1: Joliene

	PC1	PC2	PC3
(Intercept)	$-5.48\ (1.28)^{***}$	$-0.01\ (1.27)$	$-0.83\ (0.65)$
accentLate Peak	$10.82\ (2.17)^{***}$	$0.07\ (2.06)$	$1.69\ (0.92)$

$^{***}p < 0.001; {}^{**}p < 0.01; {}^{*}p < 0.05$

Table 4.4: Statistical Models Experiment 1.1: Marjolien

	PC1	PC2	PC3
(Intercept)	$0.21\ (0.92)$	$-2.87\ (0.64)^{***}$	$0.01\ (0.31)$
accentLate Peak	$-0.43\ (2.27)$	$5.48\ (0.82)^{***}$	$0.01\ (0.59)$

$^{***}p < 0.001; {}^{**}p < 0.01; {}^{*}p < 0.05$

Table 4.5: Statistical Models Experiment 1.2: Lin

	PC1	PC2	PC3
(Intercept)	-0.87 (1.33)	-3.33 $(0.82)^{***}$	1.33 $(0.46)^{**}$
accentLate Peak	1.89 (2.49)	6.92 $(0.97)^{***}$	-2.74 $(0.52)^{***}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 4.6: Statistical Models Experiment 1.2: Lina

	PC1	PC2	PC3
(Intercept)	9.00 $(1.92)^{***}$	-2.07 $(1.05)^{*}$	0.23 (0.43)
accentLate Peak	-18.15 $(2.84)^{***}$	4.05 $(1.41)^{**}$	-0.49 (0.91)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

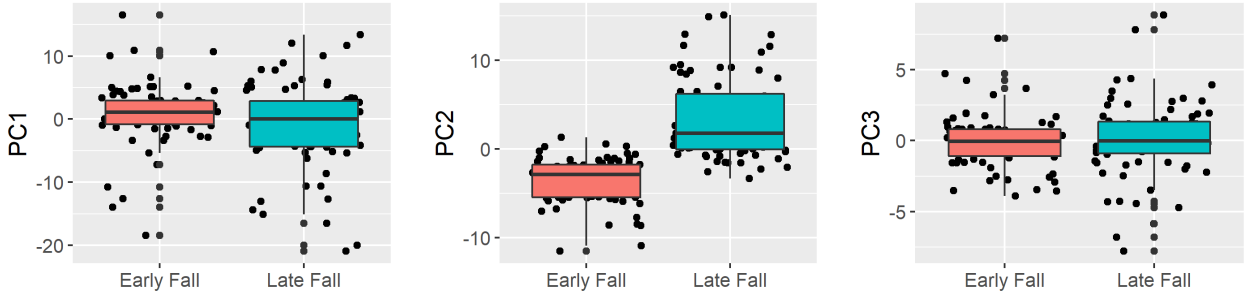
Table 4.7: Statistical Models Experiment 1.2: Alina

	PC1	PC2	PC3
(Intercept)	8.59 $(1.73)^{***}$	-0.73 (1.04)	-2.23 $(0.81)^{**}$
accentLate Peak	-11.47 $(3.22)^{***}$	1.55 (2.61)	3.76 $(1.03)^{***}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

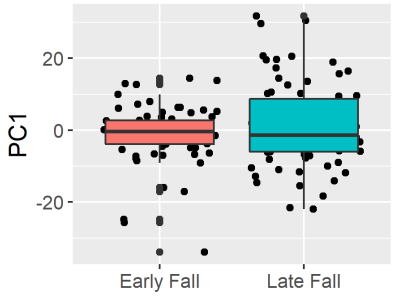
Table 4.8: Statistical Models Experiment 1.2: Evelin

Appendix B: Principle Component Graphs and Trajectories

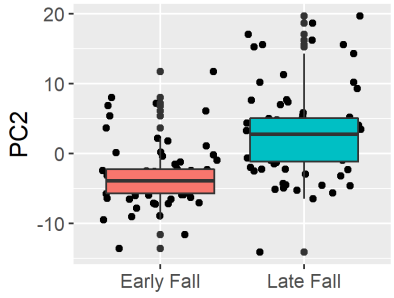


(a) Principle Component 1 *Lien* (b) Principle Component 2 *Lien* (c) Principle Component 3 *Lien*

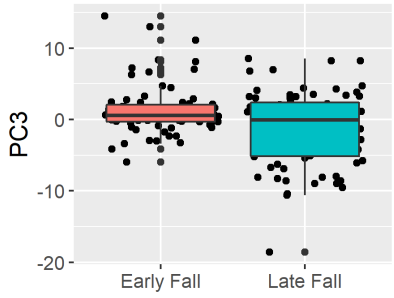
Fig. 4.14: PCA Experiment 1.1



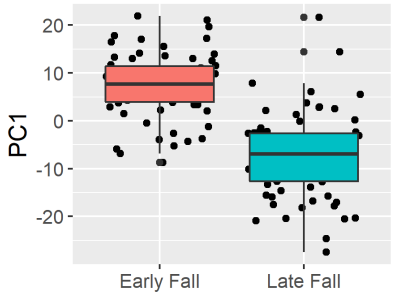
(a) Principle Component 1 *Lieneke*



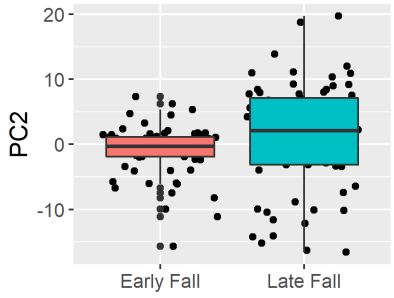
(b) Principle Component 2 *Lieneke*



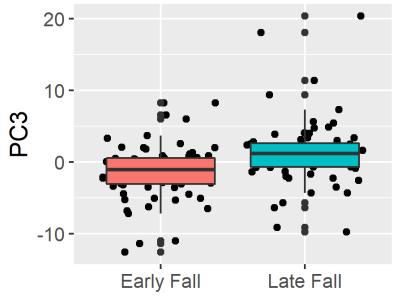
(c) Principle Component 3 *Lieneke*



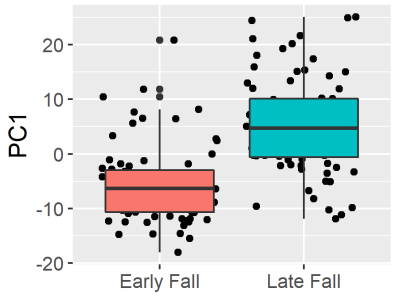
(d) Principle Component 1 *Joliene*



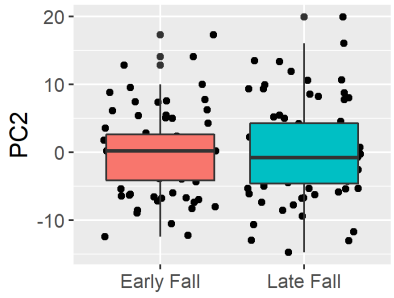
(e) Principle Component 2 *Joliene*



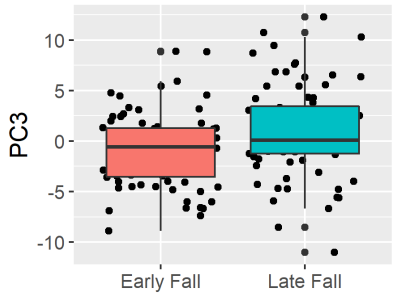
(f) Principle Component 3 *Joliene*



(g) Principle Component 1 *Marjolien*

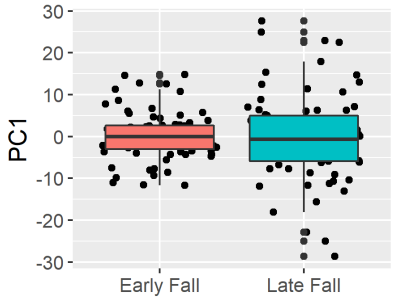


(h) Principle Component 2 *Marjolien*

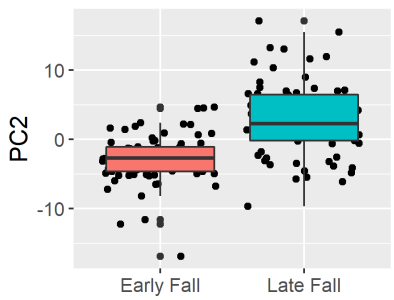


(i) Principle Component 3 *Marjolien*

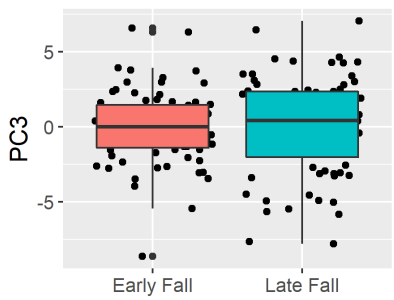
Fig. 4.15: PCA Experiment 1.1



(a) Principle Component 1 *Lin*

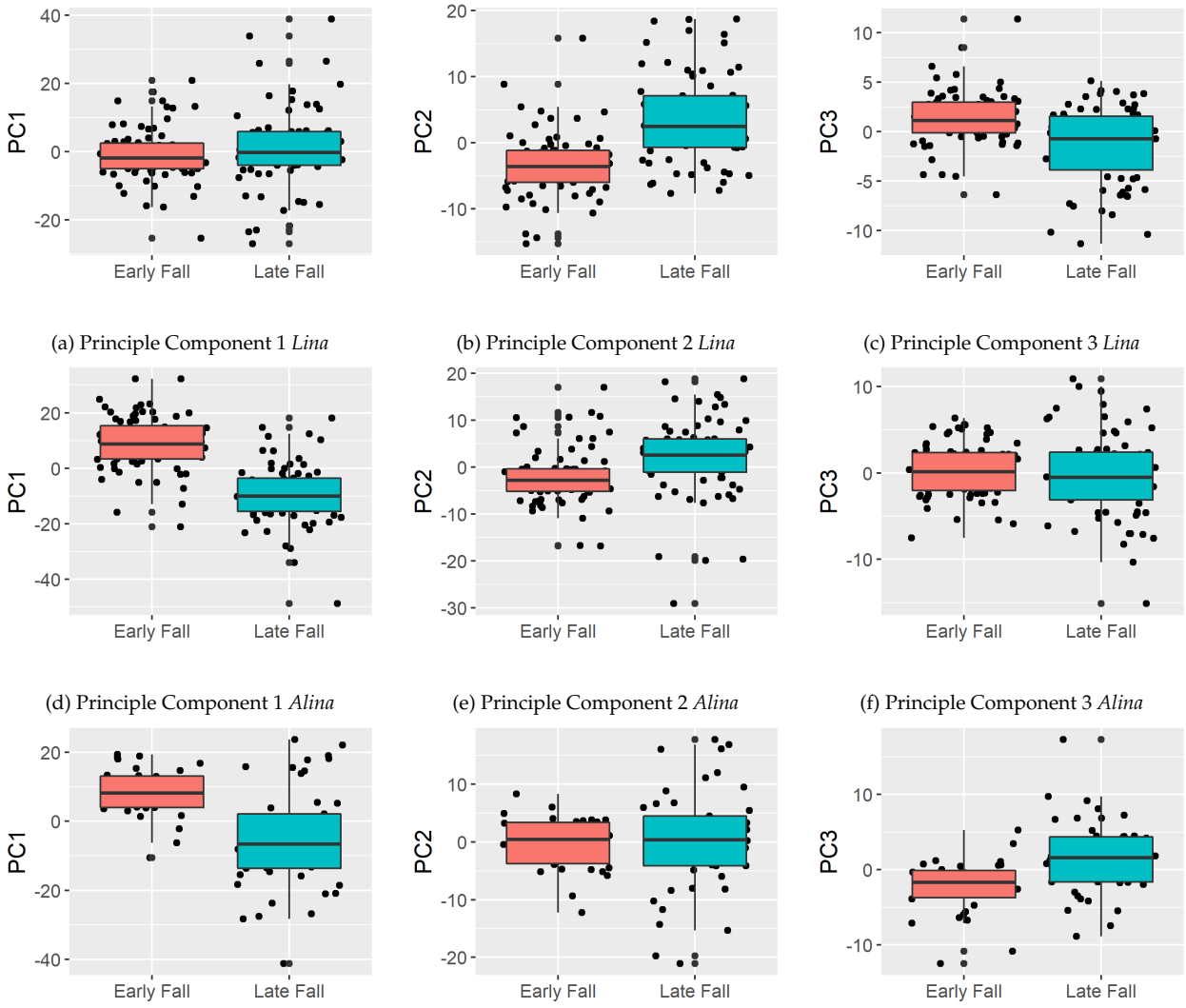


(b) Principle Component 2 *Lin*



(c) Principle Component 3 *Lin*

Fig. 4.16: PCA Experiment 1.2



(a) Principle Component 1 *Lina*

(b) Principle Component 2 *Lina*

(c) Principle Component 3 *Lina*

(d) Principle Component 1 *Alina*

(e) Principle Component 2 *Alina*

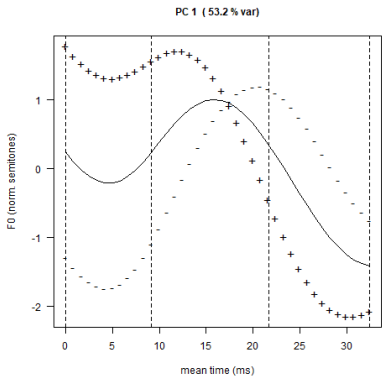
(f) Principle Component 3 *Alina*

(g) Principle Component 1 *Evelin*

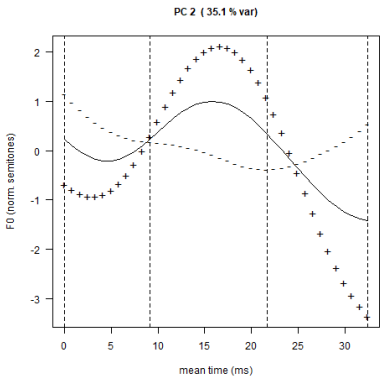
(h) Principle Component 2 *Evelin*

(i) Principle Component 3 *Evelin*

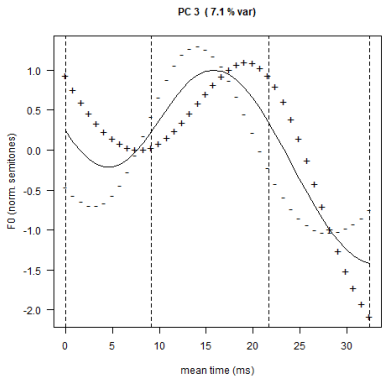
Fig. 4.17: PCA Experiment 1.2



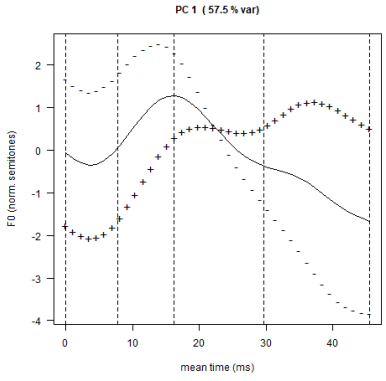
(a) Principle Component 1 *Lien*



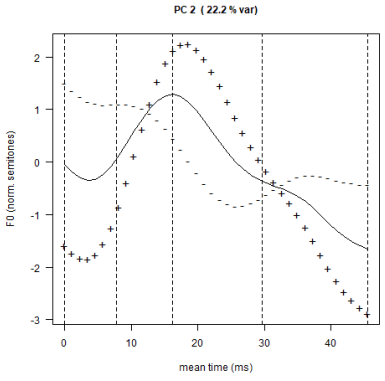
(b) Principle Component 2 *Lien*



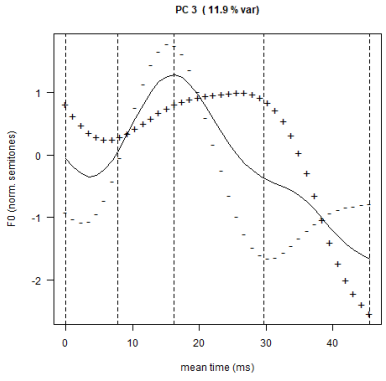
(c) Principle Component 3 *Lien*



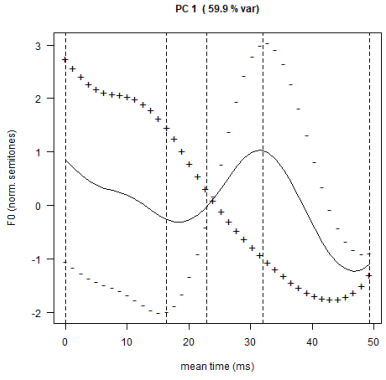
(d) Principle Component 1 *Lieneke*



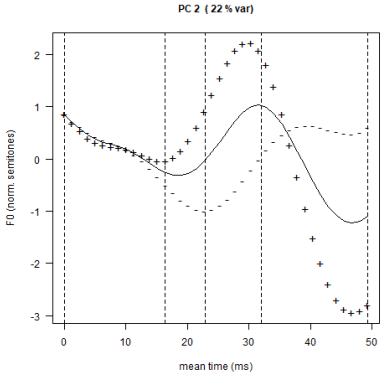
(e) Principle Component 2 *Lieneke*



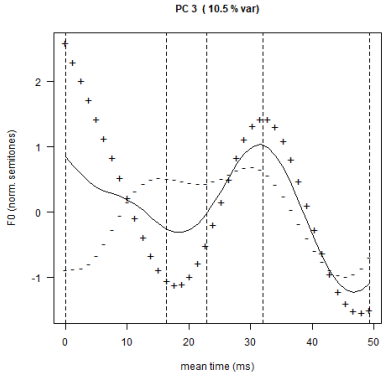
(f) Principle Component 3 *Lieneke*



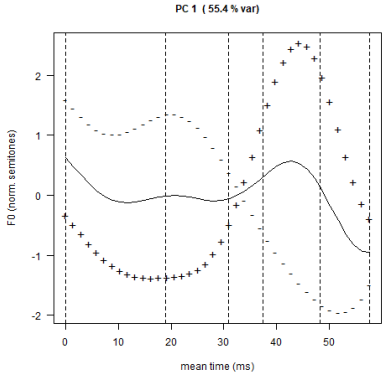
(g) Principle Component 1 *Joliene*



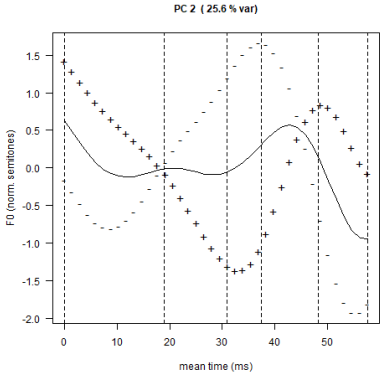
(h) Principle Component 2 *Joliene*



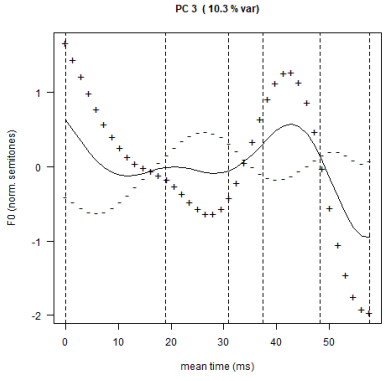
(i) Principle Component 3 *Joliene*



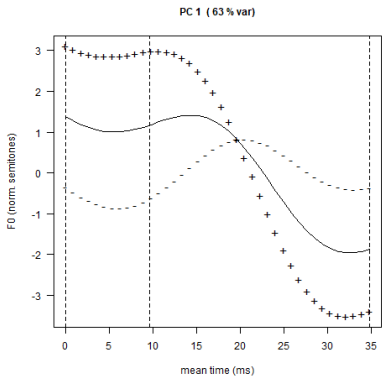
(j) Principle Component 1 *Marjolien*



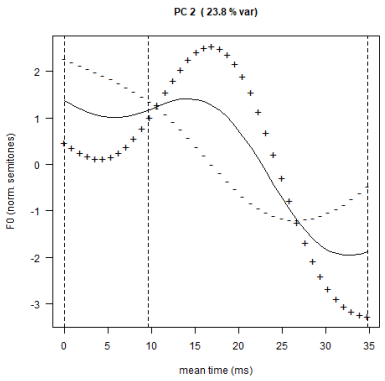
(k) Principle Component 2 *Marjolien*



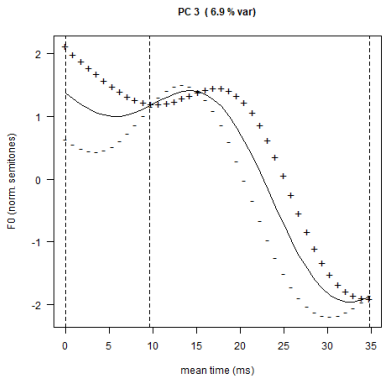
(l) Principle Component 3 *Marjolien*



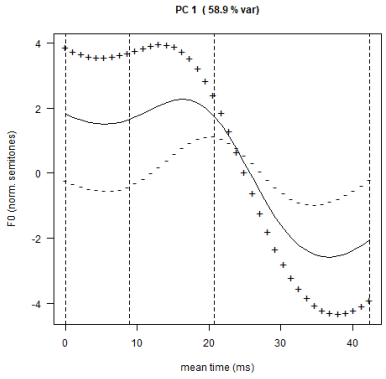
(a) Principle Component 1 *Lin*



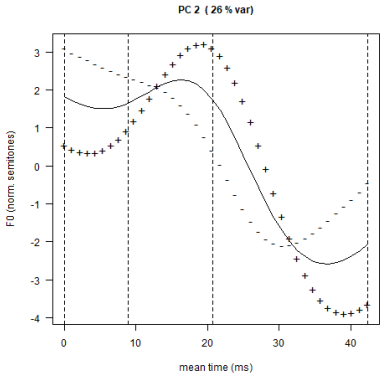
(b) Principle Component 2 *Lin*



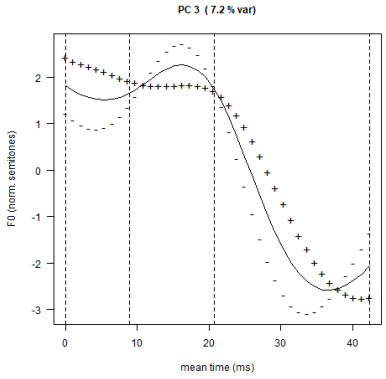
(c) Principle Component 3 *Lin*



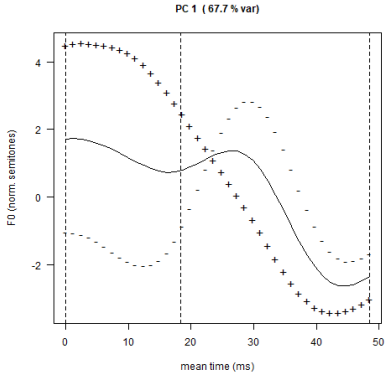
(d) Principle Component 1 *Lina*



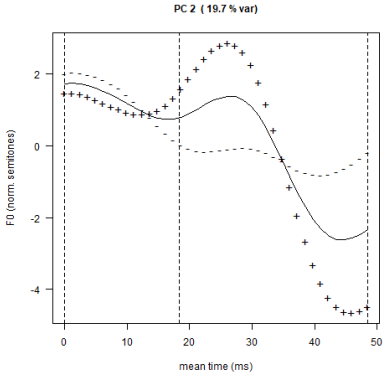
(e) Principle Component 2 *Lina*



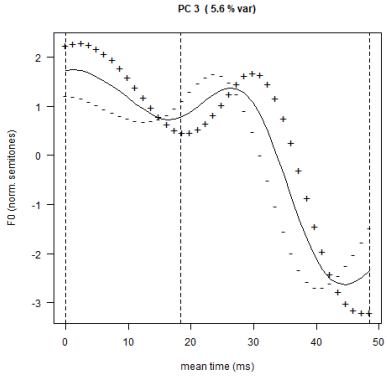
(f) Principle Component 3 *Lina*



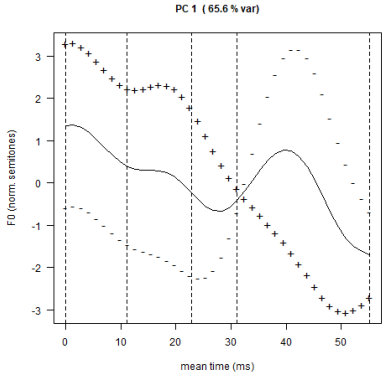
(g) Principle Component 1 *Alina*



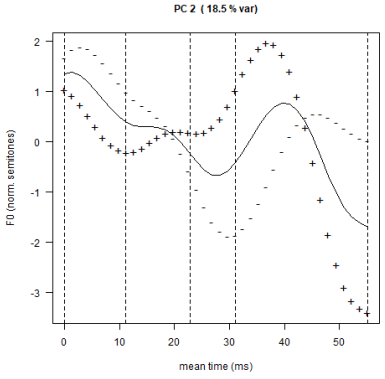
(h) Principle Component 2 *Alina*



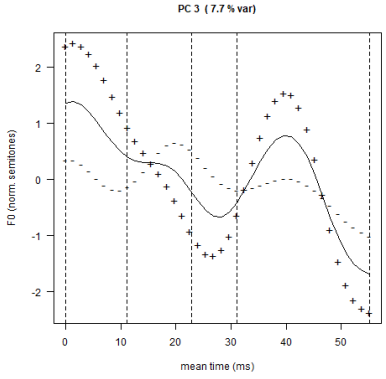
(i) Principle Component 3 *Alina*



(j) Principle Component 1 *Evelin*



(k) Principle Component 2 *Evelin*



(l) Principle Component 3 *Evelin*

Chapter 5

Topic-Focus Constructions in German and Dutch

The previous chapter explored the phonological differences between the early and late peak accents in German and Dutch. These two accents are said to be crucial to the meanings ascribed to hat patterns, more about the semantics of early and late peak hat patterns will follow in the next two chapters. This chapter explores sentences consisting of a contrastive topic and focus in German and Dutch, and the intonation patterns that go with such sentences. Elaborating on Jackendoff (1972) and combining it with Rooth (1985)'s ideas of alternative semantics, Büring (1997) provides a formal analysis on the semantics of the hat pattern in German. Büring (1997) follows Jackendoff (1972)'s topic-focus analysis of B and A accents and applies a similar analysis to the hat pattern in German: The basic idea is that contrastive topics are marked by the rising gesture of the hat pattern and foci by the falling gesture of the hat pattern. This chapter tests the claims made by Büring (1997) with regards to the hat pattern. Similar to English, Büring (1997)'s account of the hat pattern is then compared to a simpler account in which prosody does not phonologically distinguish contrastive topics and contrastive foci, rather, non-final foci can optionally be

realized in a way that corresponds to what other accounts have argued are contrastive topic realizations.

5.1 Hat pattern as Topic-Focus Construction

Büring (1997) takes the hat pattern (or the bridge pattern as he calls it) to be composed of two phonological units: the initial rise and the final fall. The rising accent is represented as L+H* (called *early rise* in this study) and the falling accent is represented as H+L* (called *early peak* accent in this study) by Büring (1997; 1999).[1] For Büring (1997) each phonological element has its own pragmatic/semantic function, the rise constitutes the sentence internal topic (S-topic) and the fall constitutes the focus of the sentence. Whichever element in the sentence receives focus will be part of the set of propositions that are considered well-formed alternatives, this Büring calls the *focus value*. If focus is assigned to a constituent X, the focus value of that sentence is the set of propositions one gets by sticking in alternatives for the focus-marked constituent.

Let us look at an example to make things clearer. A wh-question in which the object is questioned (see Example 53) introduces a focused constituent in the answer. The focus value of the answer consists of a set of propositions in which the focused element (in this case the object) can be substituted with any alternative *x* (see 54a). Answer 1 provides an example of a possible answer, in this answer *apples* receives focus and is therefore marked by a falling accent according to Büring (1997). The focus value in this specific example could be a set like this: {*The chefs ate apples, the chefs ate pears, the chefs ate pies...etc.*}.

(53) What did the chefs eat?

(54) a. Answer: The chefs ate *X*

b. Answer 1: The chefs ate APPLES \

[1] Büring (1997; 1999) provides few details on the actual phonological form of the rising and falling accents but presumably these representations correspond to what is referred to in the current study as the *early rise* and *early peak* accents respectively

S-topics (or contrastive topics) are similar to focus in the sense that they both induce alternatives. These alternatives however, do not have any impact on the focus value, instead it is a set containing different focus values, i.e. a set of sets of propositions, this Büring calls the *topic value*. The subject *waiters* in Example (55) for instance, is such a contrastive topic: Answer 2 is an element in the set of sets which make up the topic value. The topic value is a set inside of which focus values are nested, one such possible topic value for Example (55) could be: { {*The chefs ate apples, the chefs ate pears, the chefs ate pies...etc.*} {*The waiters ate apples, the waiters ate pears, the waiters ate pies...etc.*} }. It should become clear that the focus value is nested within the Topic value and by changing the subject, the topic of the sentence will change. According to Büring (1997; 1999), the topic is indicated by the initial rise in a hat pattern.

(55) Answer 2: The /WAITERS ate PEARS \
Residual Topic: What did the chefs eat?

Notice how Answer 2 does not directly answer the question *"What did the chefs eat?"*, rather it introduces a new Topic: "As for the waiters, they ate pears". The reason that a question like in (53) can be answered with an indirect answer like (55) is due to a certain implicature carried by contrastive topics, namely a disputability implicature. This implicature states that there is a question in the set of questions (i.e. topic value) which is still disputable. A question is said to be disputable if is is unresolved and there are informative but non-absurd answers to it. The question in (53) is an example of such a disputable question in the context of an answer like (55): Answer 2 does not directly resolve the question *"What did the chefs eat"* and this question therefore still stays disputed.

Even in cases where there is a direct answer, the contrastive topic implicates the existence of a disputability, i.e. an unresolved question. If, for instance one answered the question *"What did the chefs eat?"* with an answer like in Example (56) instead of Answer 1 (54b), it is argued that there is still some disputability even though the question has been answered. An answer like Answer 3 could be used to imply that there is at least one other

person/group of people who ate something besides the chefs (e.g. the waiters). According to Büring (1997) this implicature is not present in an answer like in Example (54b) which does not contain a contrastive topic.

(56) Answer 3: The /CHEFS ate APPLES \

So far Büring (1997)'s analysis of German is very similar to Jackendoff (1972)'s analysis of English, however besides the disputability implicature, one other major difference is that a contrastive topic in German must be followed by at least one focus. Therefore there are no sentences in which the (contrastive) topic follows the focus and as a consequence there are also no intonation patterns with a sequence of a falling accent followed by a rising accent. Féry (1993) argues that this is due to a phonological constraint that does not allow for a rise to follow a fall in German.

5.2 Multiple Focus Analysis

The idea that the hat pattern is composed of a Topic and a Focus element has long been the general accepted norm, but an increasing number of studies have argued against a topic and focus structure, and favour a multiple focus structure analysis (Van Hoof 2003; Ludwig 2006; Wagner 2008; 2012).

The simplest way of implementing such a multiple focus account is to treat sentences with contrastive topics as cases of multiple focus: As we have seen for English in Chapter 2, the simplest implementation of such an account is referred to as the Simple-Focus account. The Simple-Focus account predicts that to the extent that both constituents are focused and hence accented, sentences in contexts giving rise to contrastive topics, and sentences in contexts with a multiple wh-question should be similar in terms of their phonetic realization. The general idea is that a non-final focus can be optionally realized with a rising accent. Under the Simple-Focus account, there is no such thing as prosodic contrastive topic marking. However, that does not mean that from an intuitive or functional

way, one of the constituents does not function as the topic of the sentence and the other as the focus. It simply means that this distinction, to the extent that it is grammatically represented at all, does not affect prosody.

In terms of the semantics conveyed by the hat pattern, for Ludwig (2006), the German hat pattern means something like "*there is at least one true proposition (sentence) that is the result of replacing both foci with respective alternatives*". Elaborating on Ludwig (2006), Wagner (2008; 2012) argues that the hat pattern presupposes that an alternative proposition is true. More specifically, it is only the hat pattern with an early peak (EP) accent, as opposed to a late peak (LP) accent, that conveys that a contextually alternative proposition is true besides the proposition expressed bearing the hat pattern. Moreover, although Büring only considers intonation patterns starting with an early rise accent as hat patterns, for Wagner the hat pattern can either start with a rise or early rise accent. The next chapter discusses Wagner (2008)'s claims in more detail, this chapter focuses on Büring (1997) and the claims he makes with regard to the hat pattern. Büring (1997)'s account of the hat pattern is compared to the Simple-Focus account in which prosody does not phonologically distinguish contrastive topics and contrastive foci.

5.3 Expectations

The hypothesis that will be tested in this chapter is Büring (1997)'s idea that the rising accent in a hat pattern marks a contrastive topic and the falling accent marks its focus. The experiment in this chapter is analogous to the one performed in Chapter 2 on English. Like Experiment 1.2 in Chapter 2 there will be three focus conditions: *First DP focus*, *Second DP focus* and *Double focus*. Based on Büring (1997), it is expected that there will be many rising accents on the first DP and falling accents on the second DP in the *Second DP focus* condition, i.e. cases in which the first DP is the topic and the second DP is the focus,

and vice versa for the *First DP focus* condition.[2] For the *Double focus* condition it is not completely clear whether there would be any topics in the sentence but one could argue that either one of the two DPs would be interpreted as a topic. Therefore one expects either a hat pattern or if both DPs are interpreted as focus a late peak accent (LP) on both, as would be expected based on a multiple wh-question. Recall also that according to both Wagner (2008; 2012) and Büring (1999), the semantics associated with the hat pattern should only be ascribed to hat patterns with early peaks (EP) as opposed to late peak accents (LP).

Wagner (2012) does not relate the hat pattern to contrastive topics specifically but relates it to the existence of alternative propositions instead. The expectation would therefore be that the early peak hat pattern is compatible with all conditions, given that in all conditions there are potentially alternative propositions. Since Wagner (2012) makes no specific predictions with regards to contrastive topics, a contour with a late peak on both constituents as well as a late peak hat pattern should be equally acceptable. The Simple-Focus account makes the exact same predictions as Wagner (2012). Both multiple focus accounts expect non-final foci to be optionally realized with a rising accent.

Two similar experiments were performed, one for Dutch (Experiment 1.1) and one for German (Experiment 1.2): The reason for performing an experiment on German is obvious given that most of the literature on hat patterns talks about German. The reason for including Dutch is first and foremost because this is the author's first language and intuitively the hat pattern in both languages seems to convey a similar meaning.

	Double focus	**First DP focus**	**Second DP focus**
Büring (1997)	EP Hat/LP-LP	LP-LP	EP Hat
Simple-Focus/Wagner	(EP) Hat/LP-LP	(EP) Hat/LP-LP	(EP) Hat/LP-LP

Table 5.1: Expectations Büring (1997) and Wagner (2012)

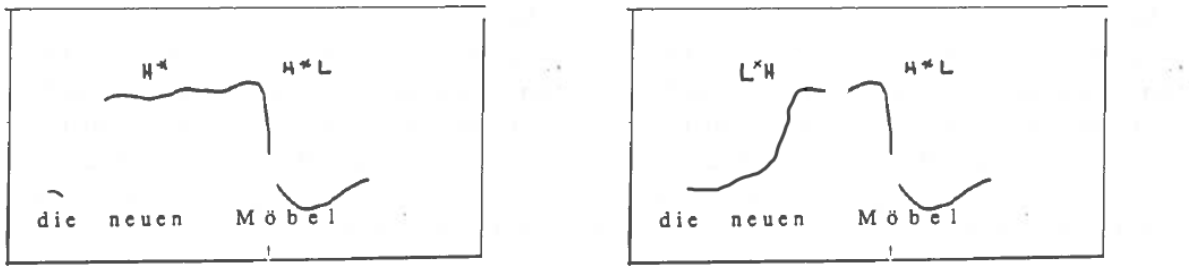
[2]Here the assumption is made that both German and Dutch would be treated similarly.

5.4 Phonology of hat patterns

Roughly speaking the hat pattern is said to start with a rise, continues as a level plateau and eventually falls in pitch thus forming an apparent hat pattern in the pitch trajectory. Despite this seemingly straightforward characterisation of the hat pattern, there is still little consensus on the correct phonological representation and there is no uniform representation to this day. Moreover, it is not clear how many types of hat patterns there are. This section provides a review of the previous literature on the phonology of the hat pattern in German and Dutch. More specifically we will discuss the pitch accents that make up a hat pattern.

5.5 Rise and Early Rise hat pattern

In her book *Prosodic and tonal structure of standard German*, Féry (1989) recognizes two different hat patterns. One consists of an initial high target tone H* (in this study referred to as early rise accent) followed by a falling accent H*L (in this study referred to as the late peak accent), as illustrated in Figure (5.1a), and another one starts off with a rising accent (L*H) followed by a falling accent (H*L), as illustrated in Figure (5.1b)[3].



(a) hat pattern 1. H*-H*L contour 'The new furniture'

(b) hat pattern 2. L*H-H*L contour 'The new furniture'

Fig. 5.1: Ferry (1989) hat pattern example

[3]In contrast to Figure (a) this Figure has been constructed based on Figure (a) for explanatory purposes

Pragmatically and phonologically, these two contours are argued to be fundamentally different. The Early Rise (H*-H*L) hat pattern (referred to as hat pattern 1) is a sequence of two fully linked pitch accents and crucially consists of only one intonational phrase according to Féry (1989). The hat pattern starting with a rise (referred to as hat pattern 2) on the other hand, consists of two intonational phrases that are linked syntactically or semantically. To date there is however no empirical evidence to support the claim that hat pattern 1 consists of just one intonational phrase and hat pattern 2 of two. The patterns produced in this experiment show us that there is no consistent correspondence between the type of hat pattern and the number of IPs, instead it seems to be random (or perhaps phonetically motivated, e.g. if there is little segmental material between two pitch accents they will be produced within the same IP, but with longer intervening material the two pitch accents will be produced in separate IPs).

Besides a difference in their phonological form and in phrasing, the two hat patterns are said to differ in meaning as well. Whereas the rise in hat pattern 2 is associated with topicality, the high plateau of hat pattern 1 is not associated with any special meaning and can be used when no emphasis is needed. The contour in Figure 5.1b, for instance, could be used to express something like *with regard to the NEW furniture....* The contour in Figure 5.1a on the other hand, does not carry this meaning according to Féry (1989). The idea that the rise in a hat pattern or the hat pattern in general, is indicative of some kind of topicality has been acknowledged by many other researchers like Jacobs (1982), Höhle (1991) and Büring (1997) for German and Keijsper et al. (1984) for Dutch. In this chapter the idea that there is some correlation between topicality and the rising accent will be put to the test. In the following section the two pitch accents that make up these two types of hat patterns will be described in more detail.

5.5.1 Early Rise

Despite describing German intonation, Féry (1989)'s transcriptions tend to align with ToDI(Transcription for Dutch Intonation)-conventions more closely than with GToBI(German Tones and Break Indices)-conventions. In ToDI, the *Early Rise* pitch accent is transcribed as *H** in Dutch and is defined as a pitch accent with sustained high pitch (Gussenhoven et al. 2005). GToBI (German ToBI) provides a more or less similar description for this pitch accent in German, in GToBI the *H** accent is perceived as relatively high and may be preceded by a shallow rise, see Figures 5.4 and 5.5 (Grice et al. 2007). In addition to *H**, GToBI acknowledges yet another pitch accent which is somewhat similar to the *H** accent, namely the *L H** accent. For the *L H** accent, the accented syllable is perceived as high, like the *H** accent, but unlike the *H** accent it is preceded by a syllable with a low pitch target which leads to a sharp rise in the accented syllable (see Figures 5.2 and 5.3). The peak is often late in the accented syllable (Adriaens 1991; Grabe et al. 1998).

As mentioned in Chapter 2, the distinction between H* and LH* is controversial and hard to perceive by transcribers (Calhoun 2012). Based on this discussion I have decided to assume only one type of *Early Rise* accent which would be transcribed as *H** following Féry (1989) and Gussenhoven et al. (2005).[4]

[4]More research is necessary to determine whether there is a semantic difference between these two accents.

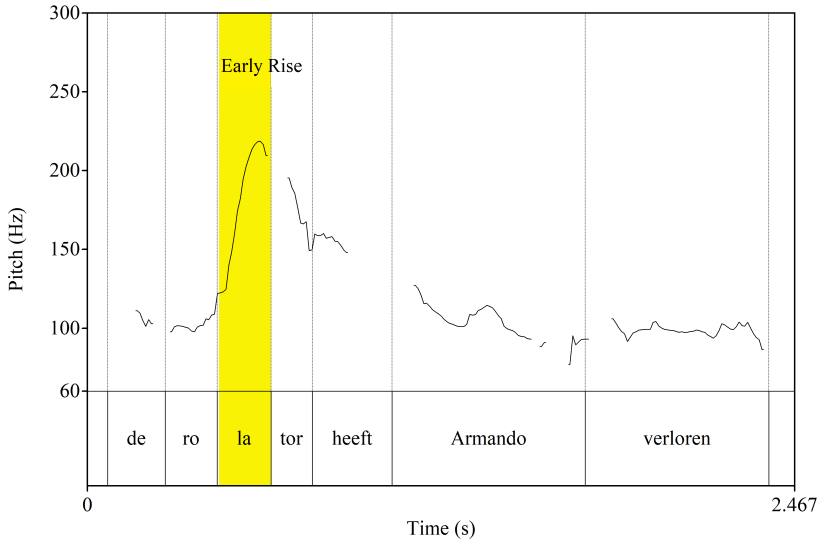


Fig. 5.2: Pitch trajectory “de rolator heeft Armando verloren" (the walker Armando lost) with a Early Rise accent on *rolator* in Dutch. The highlighted part in yellow marks the accented syllable. The pitch starts off lower than the Early Rise accent thus giving rise to a rise. Click here to listen.

Fig. 5.3: Pitch trajectory “Johann wurde von Nina kritisiert" (Johann was critized by Nina) with a Early Rise accent on *Johann* in German. The pitch starts off low thus giving rise to a rise. Click here to listen.

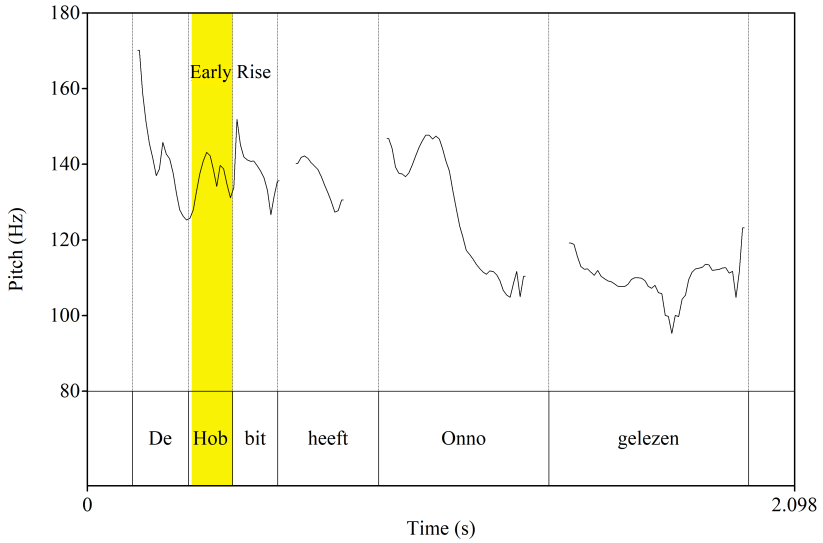


Fig. 5.4: Pitch trajectory “De Hobbit heeft Onno gelezen" (The Hobbit Onno read) with a Early Rise accent on *Hobbit* in Dutch. Click here to listen.

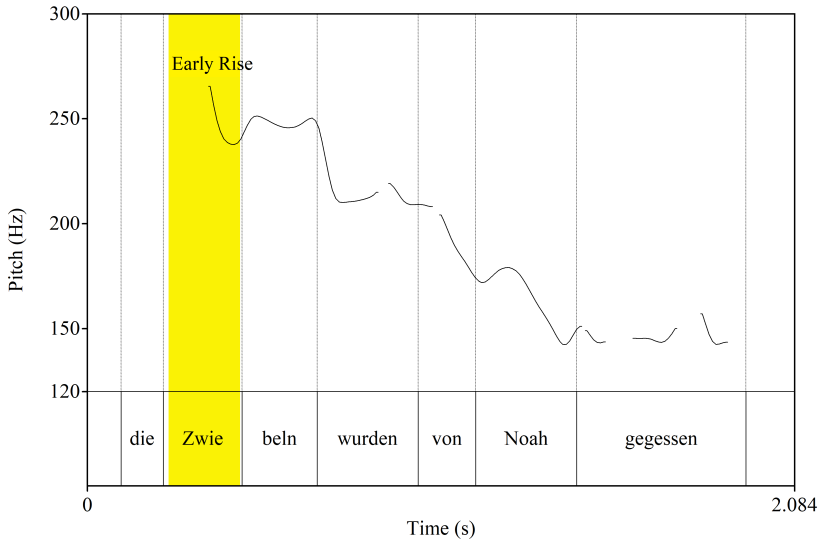


Fig. 5.5: Pitch trajectory “die Zwiebeln wurden von Noah gegessen" (the onions Noah ate) with a Early Rise accent on *Zwiebeln* in German. Click here to listen.

5.5.2 Rise accent

Different from the *Early Rise*, the (late) rising accent receives the same analysis in both ToDI and GToBI and is transcribed as L*+H (i.e. a low target tone followed by a high trailing tone): There is a low target within the accented syllable which is followed by a rise, starting late in the accented syllable and reaching its peak on the next syllable (or sometimes later). In contrast to H*, the perceived pitch of the accented syllable is low (see Figures 5.6 and 5.7).

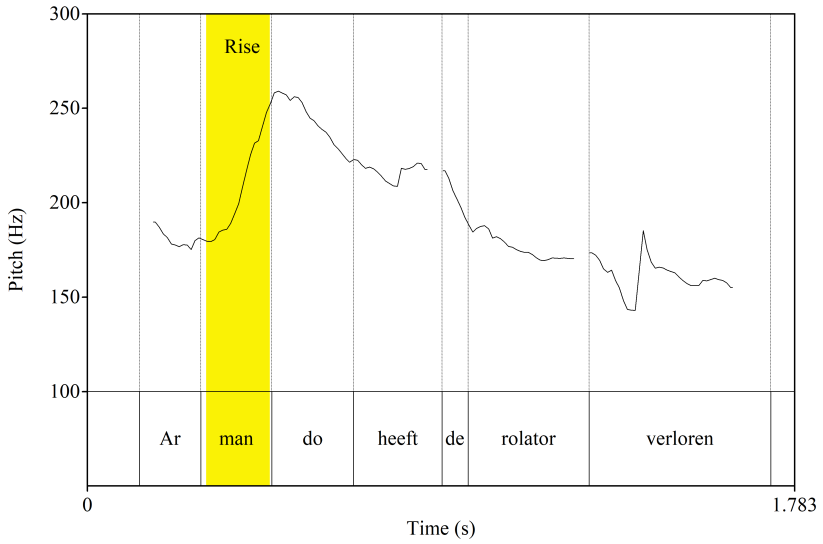


Fig. 5.6: Pitch trajectory “Armando heeft de rolator verloren" (Armando lost the walker) with a rising accent on *Armando* in Dutch.Click here to listen.

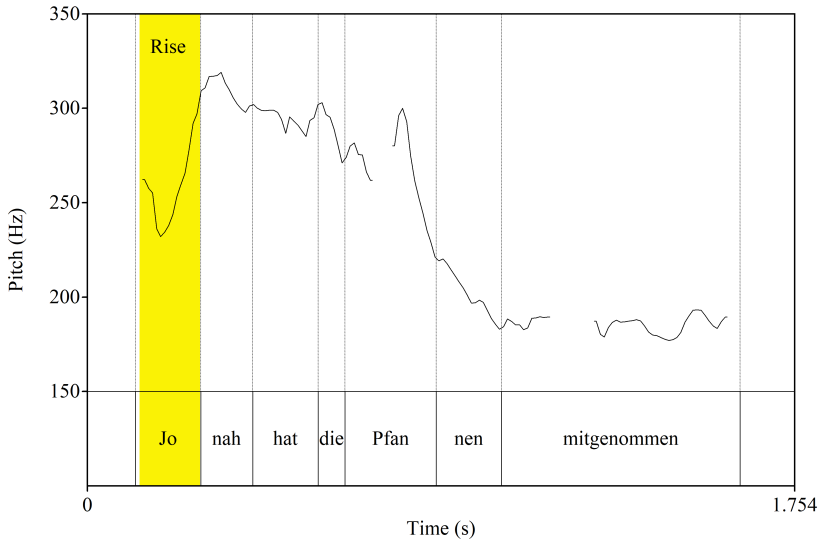


Fig. 5.7: Pitch trajectory “Jonah hat die Pfannen mitgenommen" (Johan brought the pans) with a rising accent on *Jonah* in German. Click here to listen.

5.6 Late Peak (LP) and Early Peak (EP) accent

Unlike Féry (1989) who does not distinguish different hat patterns based on the falling accent, Wagner (2008; 2012) argues that the semantics assigned to the hat pattern only applies to the hat pattern with a ‘sharp fall’ whereas the other falling accent is more neutral and does not have any specific semantics assigned to it. In this study the ‘sharp fall’ is referred to as the early peak accent, as opposed to the late peak accent.[5] The label *‘early’* here is roughly used to indicate that an f0 peak precedes the accented syllable (see Figures 5.10 and 5.11) and the label *‘late’* indicates that an f0 peak occurs within the accented syllable (see Figures 5.8 and 5.9). These two pitch contours have been extensively researched in the previous chapter and the conclusion was drawn that the early peak contour is best described as (%H) H!H* (L%) and the late peak contour as H* L%.

[5]Büring (1997) represents the hat pattern with an early peak accent as well, but does not go into much detail about exact phonetic implementation.

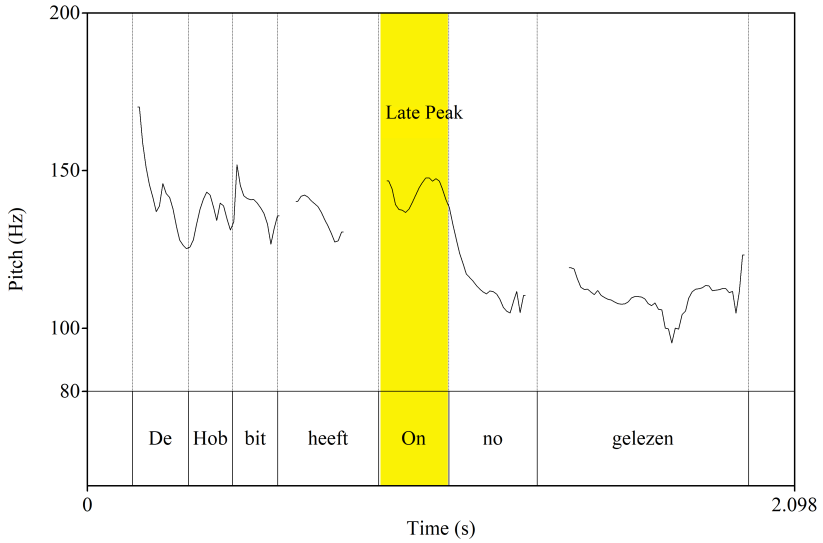


Fig. 5.8: Pitch trajectory “De Hobbit heeft Onno gelezen" (The Hobbit Onno read) with a Late Peak accent on *Onno* in Dutch. Click here to listen.

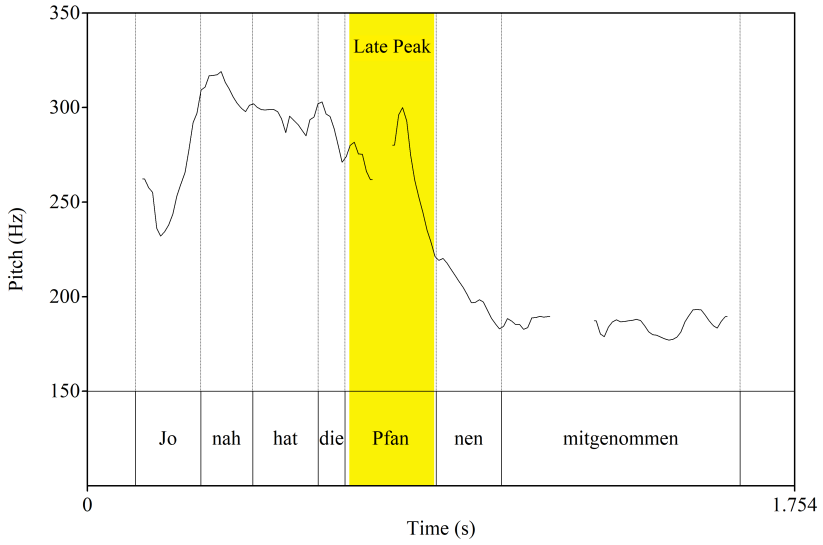


Fig. 5.9: Pitch trajectory “Jonah hat die Pfannen mitgenommen" (Johan brought the pans) with a Late Peak accent on *Pfannen* in German. Click here to listen.

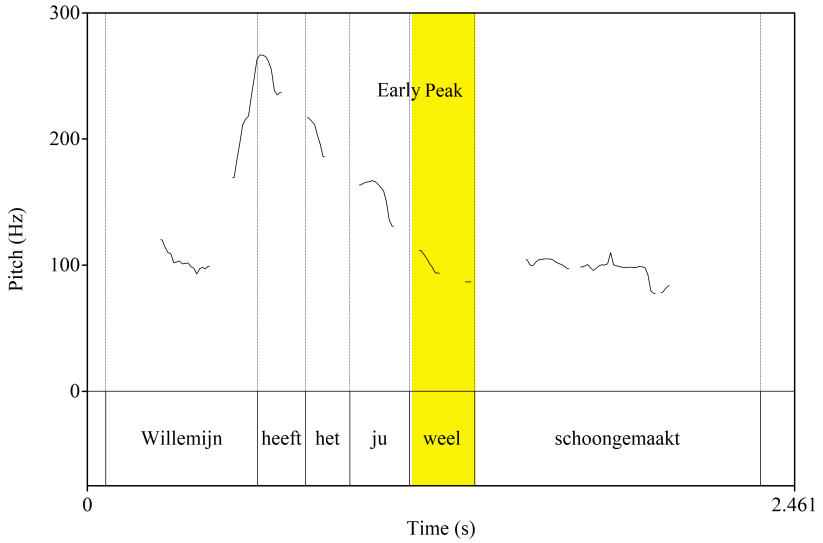


Fig. 5.10: Pitch trajectory “Willemijn heeft het juweel schoongemaakt" (Willemijn polished the jewelry) with an Early Peak accent on *juweel* in Dutch. Click here to listen.

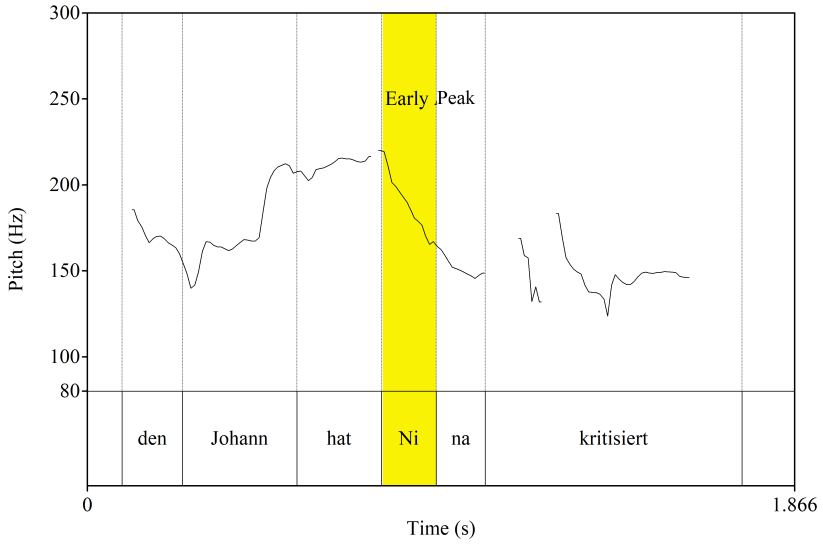


Fig. 5.11: Pitch trajectory “Den Johann hat Nina kritisiert" (John Nina criticized) with an Early Peak accent on *Nina* in German. Click here to listen.

5.7 Experiment 1.1: Dutch

In order to test the claims made by Büring (1997), an online production experiment was set up using jsPsych (de Leeuw 2015). Participants were presented with a number of carefully manipulated questions, both in written form as well as auditorily, after which they were asked to record their answer and assess how well the answer sounded.

5.7.1 Participants

A total of 30 participants were recruited through Prolific, all of them were native speakers of Standard Dutch (as spoken in the Netherlands). A compensation of 2.33 C$ was given which boils down to an hourly rate of 13.98 C$ since the experiment itself took about 10 minutes.

5.7.2 Stimuli

Each stimulus consists of a question and an answer. There are three different focus conditions and each question is preceded by a general context (see Example 57): Double focus, object focus and subject focus. The *double focus* condition consists of a multiple wh-question (see 57a), the *object focus* condition consists of a wh-question questioning the object (see 57b) and the *subject focus* condition consists of a wh-question questioning the subject (see 57c). The contexts plus questions have been recorded by the author and serve as an auditory stimuli to be presented to the participants during the experiment.

(57) *Dutch*

Ik hoorde dat enkele studenten elkaar hebben bekritiseerd.
I heard that some students each.other have criticized

'I heard the students criticized each other.'

a. *Double focus*

Wie heeft wie bekritiseerd?
who has whom criticized

'Who criticized whom?'

b. *Object focus*

Hoe zit het met Emma? Wie heeft zij bekritiseerd?
How sits it with Emma who has she criticized

'How about Emma, who did she criticized?'

c. *Subject focus*

Hoe zit het met Johan? Wie heeft hem bekritiseerd?
How sits it with John who has him criticized

'How about John? Who criticized him?'

For the answers there three different word orders: Active passive and topic syntax (see Example 58). What is crucial for these sentences is the order of the object and subject of the sentence.

(58) a. *Active syntax*

Emma heeft Johan bekritiseerd.
Emma has John criticized.

'Emma has criticized John.'

b. *Passive syntax*

Johan werd door Emma bekritiseerd.
John was by Emma criticized.

'John was criticized by Emma.'

c. *Topicalized syntax*

Johan heeft Emma bekritiseerd.
John has Emma criticized.

'John, Emma criticized.'

In total there are nine conditions, i.e. three questions conditions and three answer conditions. 36 stimuli sets were created each condition would be repeated four times.

For analysis purposes the focus conditions have been relabeled to *Double focus*, *First DP focus* and *Second DP focus*. The first DP is focused in the object focus condition for sentences with a passive and topicalized syntax, and in the subject focus condition for sentences with an active syntax (*First DP Focus*). The opposite is true for *Second DP focus*: The second DP is focused in the subject focus condition for sentences with a passive and topicalized syntax, and in the object focus condition for sentences with an active syntax (*Second DP focus*). In total there are nine conditions, i.e. three questions conditions and three answer conditions. 36 stimuli sets were created such that each condition would be repeated four times.

5.7.3 Procedure

Each participant had to perform 36 trials which were presented in a latin-square design. Participants are directed to the experiment via an online link. Each trial starts off with a fixation point, then, the context and question are given in written form and finally on the next screen, the context and question are presented auditorily. Each question ends with a beep after which the participant is asked to record the question as natural as possible. The trial finishes off with an acceptability input screen in which the participant has to indicate how natural their answer sounded, based on a scale from 1 to 8 where 1 means completely unnatural and 8 means completely natural.

5.7.4 Results

A total of 1080 recordings were made (30 participants*36 trials = 1080), one participant had to be excluded since this person was non-native and an additional 14 recordings had to be excluded due to poor sound quality, leaving us with 1030 good recordings. Eventually, 735 out of 1030 recordings contained no deaccentuations and were used for data analysis. For the results the recorded sentences were annotated by the author in terms of the pitch accents on the first DP and second DP. The following categories were distinguished for

the annotation: Late Peak, Early Peak, Early Rise, Rise, Rise-Fall-Rises[6] and Deaccented. For a phonological description of each pitch accent see the previous section.

Figure 5.12 shows the proportions of pitch accents in percentage for each condition. There is noticeably more deaccentuation on the second DP when the first DP is focused, but the inverse is not observed, i.e. more deaccentuation on the first DP when the second DP is focused. Büring (1997) is mainly interested in cases where both DPs are accented, i.e. without deaccentuation, therefore Figure 5.13 where all cases with deaccentuation have been excluded, is more informative: the exact numbers can be found in the Appendix.

Figure 5.14a gives a visualization of the distribution of the late peak accent across different conditions. Overall, it seems there are more late peaks on the second DP than on the first and the *First DP Focus* condition has the highest proportion of late peak accents. A poisson regression was run with as independent variable the proportion of late peak accents and as dependent variables the *focus*, *syntax* and *linear order* conditions as well as the interaction between the variables.[7] The final model performed significantly better than an intercept-only base line model (χ^2(13): 59.2, p = 0) and is a good, not optimal fit (C: 0.7666042, Somers' Dxy: 0.5332084). The model's explanatory power is weak (Nagelkerke's R2 = 0.12).[8] From the final model it becomes clear that only *linear order* has a significantly effect: More Late Peak accents were produced on the second DP than on the first DP (beta = 0.47, 95% CI [0.21, 0.73], p < .001).

The distribution of the Early Peak accents across the different conditions is shown in Figure 5.14b. No clear pattern can be observed from this figure other than that no early peaks are produced on the first DP. A poisson regression was run with as independent variable the proportion of early peak accents and as dependent variables *focus*, *syntax* and *linear order*. The final model performed significantly better than an intercept-only base

[6]The phonological description for this accent was not provided in this chapter but it can be found in chapter 2 for English. The assumption is that the RFR accents in English German and Dutch are the same.
[7]See Appendix B for the statistical models.
[8]The model's intercept, corresponding to focus = Double focus, constituent = First DP and syntax = Active, is at -0.58 (95% CI [-0.81, -0.36], p < .001).

line model (χ^2(5): 76.101, p = 0) and is a good, optimal fit (C: 0.0.8057179, Somers' Dxy: 0.6114358). The model's explanatory power is moderate (Nagelkerke's R2 = 0.24). No significant effects were found. It should also be noted that no statistical analysis could be performed on the linear order condition since there were zero productions on the first DP.

The distribution of the rising accents across the different conditions is shown in Figure 5.15a. The results tell us that more rising accents were produced on the first DP than on the second and the *First DP focus* condition seems to have a lower proportion of rising accents than the other conditions. A poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of rise accents was taken as independent variable and as dependent variables *focus*, *syntax* and *linear order* were included in the model. The final model performed significantly better than an intercept-only base line model (χ^2(5): 176.35, p = 0) and is a good, optimal fit (C: 0.8129486, Somers' Dxy: 0.6258972). The model's explanatory power is substantial (Nagelkerke's R2 = 0.32).[9] From the final model it becomes clear that *focus* has a significantly effect: Less rising accents were produced in the *First DP focus* condition than in other focus condition (see the full model in Appendix B for the estimates). Moreover, a significant effect was found for *linear order*: Less rising accents were produced on the second DP than on the first DP (beta = -4.17, 95% CI [-5.97, -3.02], p < .001).[10]

Figure 5.15b gives a visualization of the distribution of the early rises across different conditions. Similar to rising accents, the results tell us that more early rises were produced on the first DP than on the second and the *First DP focus* condition seems to have a lower proportion of early rises than the other conditions. A similar poisson regression to the one for rising accents was run. The final model performed significantly better than an intercept-only base line model (χ^2(5): 184.9, p = 0) and is a good, optimal fit (C: 0.8033638, Somers' Dxy: 0.6067276). The model's explanatory power is substantial (Nagelkerke's R2

[9] The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = First DP focus, is at -2.61 (95% CI [-3.36, -1.98], p < .001).

[10] Although no interaction effect was found, this could be due to the fact that in some conditions there were zero productions.

= 0.32).[11] A significant effect was found for *linear order*: Less early rises were produced on the second DP than on the first DP (beta = -4.26, 95% CI [-6.06, -3.12], p < .001).

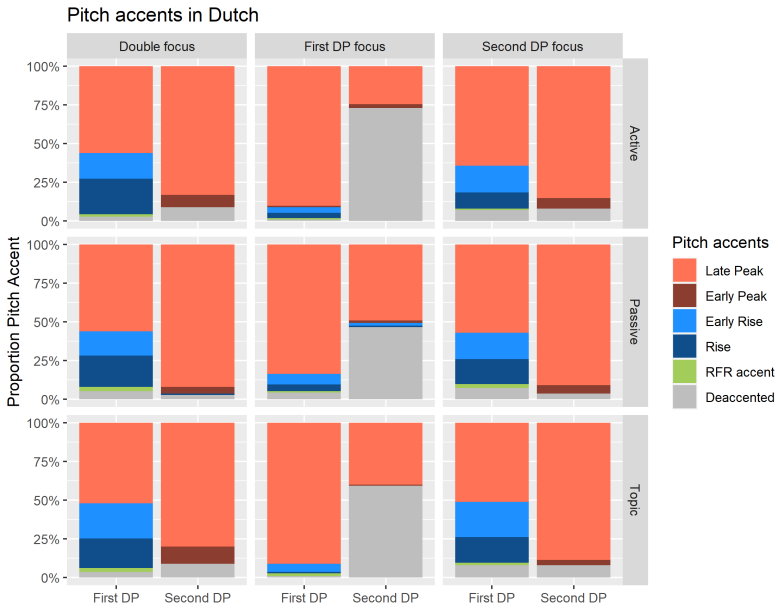


Fig. 5.12: Pitch Accents with Deaccentuation

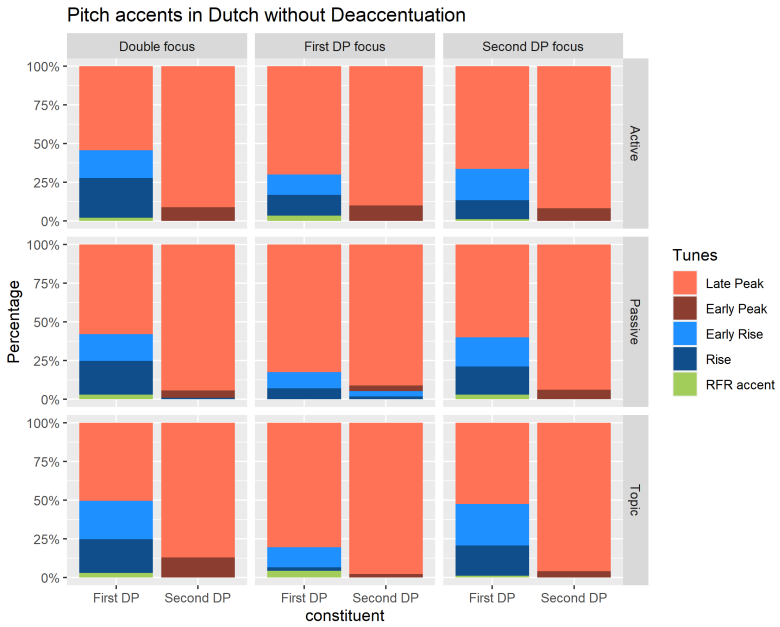
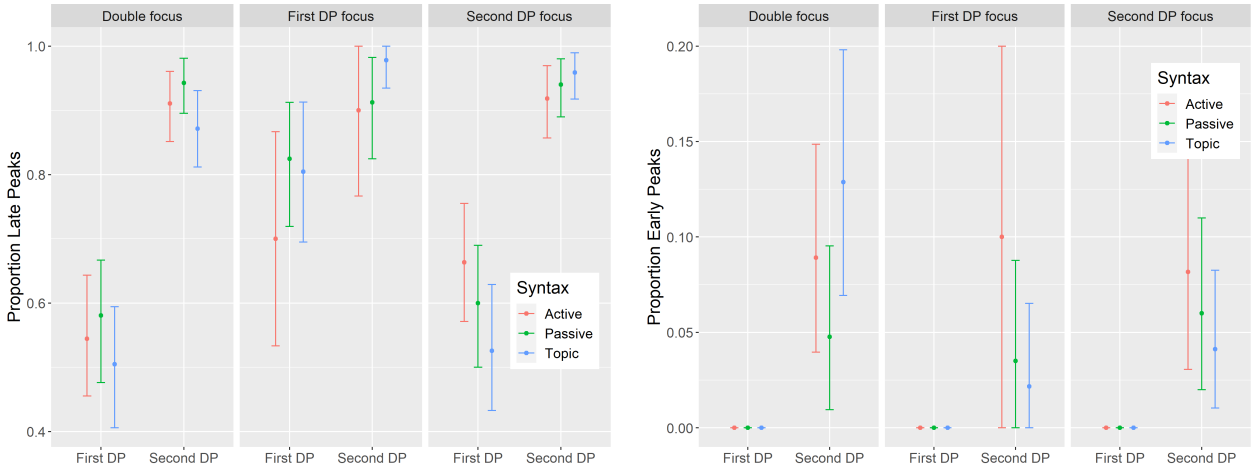


Fig. 5.13: Pitch Accents without Deaccentuation in Dutch

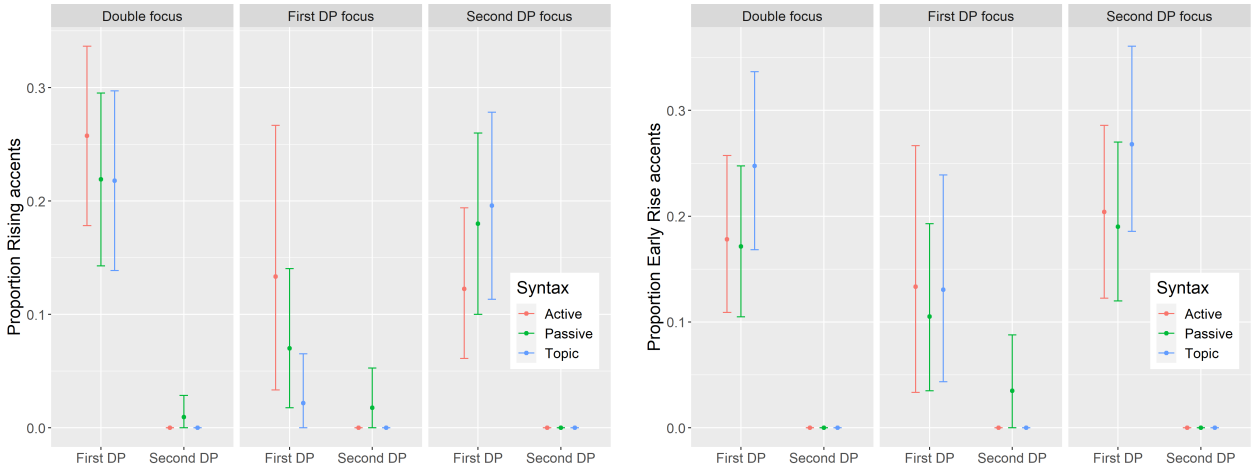
[11] The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = First DP focus, is at -2.10 (95% CI [-2.68, -1.59], p < .001).



(a) Distribution of Late Peak accents

(b) Distribution of Early Peak accents

Fig. 5.14: Distribution Pitch accents



(a) Distribution of Rising accents

(b) Distribution of Early Rising accents

Fig. 5.15: Distribution Pitch accents

5.7.5 Conclusion

The results from Experiment 1.1 seem to partially support Büring (1997)'s claim that contrastive topics are marked by a rising accent (although it is not the early rising accent that Büring presumably had in mind) and foci by a falling accent, at least for Dutch: There are significantly less rising accents in the First DP Focus condition than in the other focus conditions. However, this pattern is rather weak since there are relatively few rising accents

to begin with. On top of that a large proportion of contrastive topics have been produced with a falling accent. In conclusion, although Experiment 1.1 does seem to partially support Büring (1997)'s analysis, the mapping between a rising pitch accent and contrastive topic is much weaker than one would have expected based on Büring (1997)'s account. Interestingly, most rising accents have been produced in the double focus condition, there is no clear explanation for why this would be the case.

5.8 Experiment 1.2: German

The methodology for Experiment 1.2 is the same as in Experiment 1.1 and only those parts where the two experiments differed have been written down.

5.8.1 Participants

A total of 30 participants were recruited, all of them were native speakers of Standard German (as spoken in Germany). A compensation of 2.33 C$ was given which boils down to an hourly rate of 13.98 C$ since the experiment itself took about 10 minutes.

5.8.2 Stimuli

Each stimulus consists of a question and an answer. There are three different question conditions and each question is preceded by a general context (see Example 59): Double focus, object focus and subject focus. The *double focus* condition consists of a multiple wh-question (see 59a), the *object focus* condition consists of a wh-question questioning the object (see 59b) and the *subject focus* condition consists of a wh-question questioning the subject (see 59c). The contexts plus questions have been recorded by the researchers and serve as an auditory stimuli to be presented to the participants during the experiment.

(59) *German*

Ich habe gehört, dass die Schüler sich gegenseitig kritisiert haben.
I have heard that the students themselves each.other criticized have

'I heard the students criticized each other.'

a. *Double focus*

Wer hat wen kritisiert?
who has whom criticized

'Who criticized whom?'

b. *Object focus*

Was ist mit Nina? Wen hat sie kritisiert?
what is with nina who has she criticized

'How about Nina, who did she criticized?'

c. *Subject focus*

Was ist mit Johann? Wer hat den kritisiert?
what is with John who has him criticized

'How about John? Who criticized him?'

For the answers there are also three manipulations: Active syntax, passive and topicalized sentences. What is crucial for these sentences is the order of the object and subject of the sentence.

(60) a. *Active syntax*

Nina hat Johann kritisiert.
B: Nina has John criticized.

'Nina has criticized John.'

b. *Passive syntax*

Johann wurde von Nina kritisiert.
B: John was of Nina criticized.

'John was criticized by Nina.'

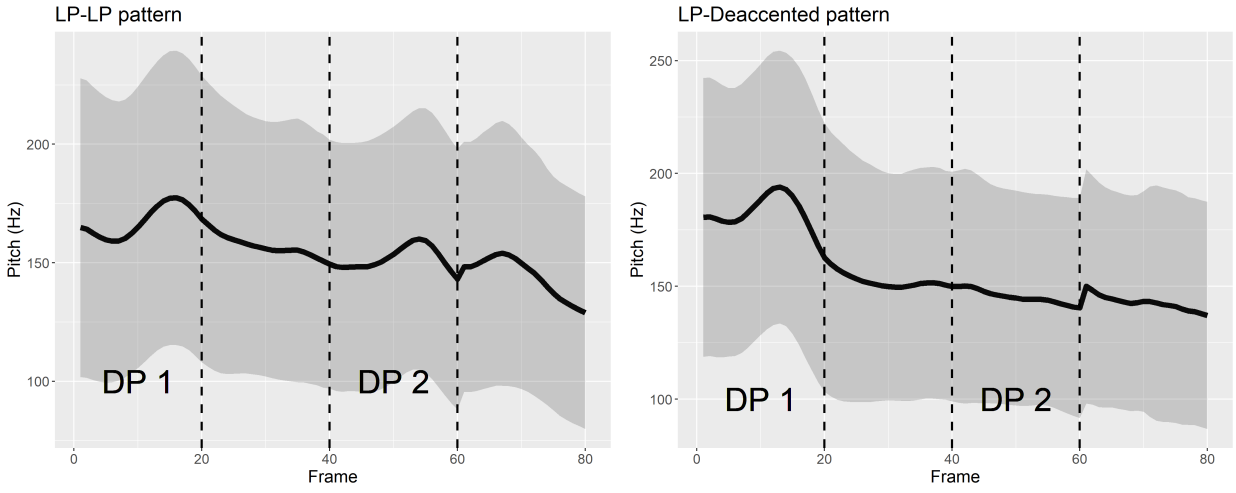
c. *Topicalized syntax*

Den Johann hat Nina kritisiert.
B: the John has Nina criticized.

'John, Nina criticized.'

5.8.3 Results: Mean Pitch Trajectory

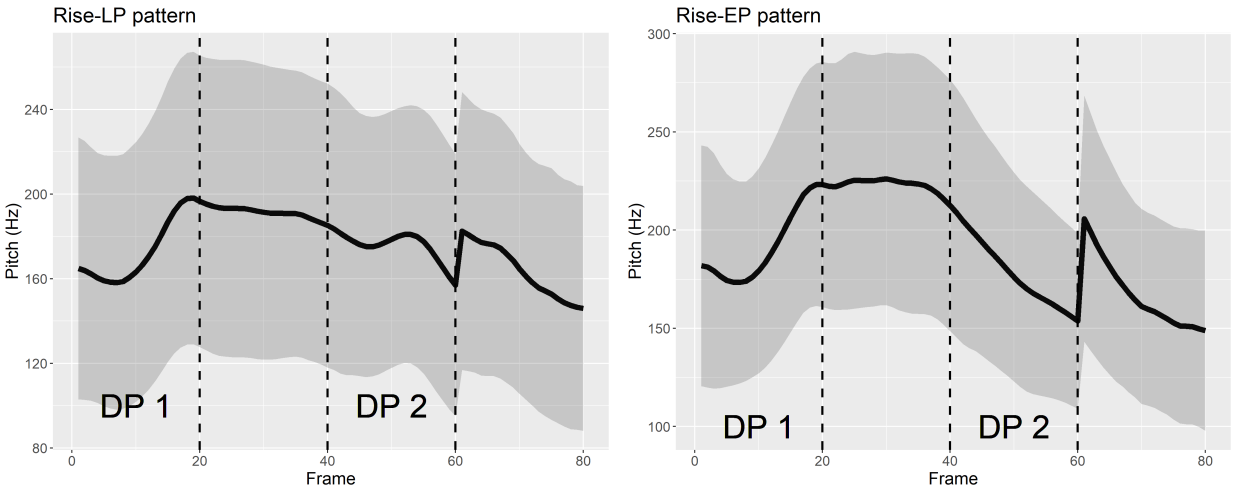
To give an idea of the actual productions, the mean pitch trajectories of some of the most frequent contours have been illustrated in this sections (see Figures 5.16 and 5.17). By far the most common contours are the LP-LP contour and LP-Deaccented contour (see Figure 5.16). The contours that are most interesting for the purpose of this research are the hat patterns, i.e. contours with a rising accent followed by either a late peak (5.17a) or early peak accent (5.17b). One clear observation that can be made from the figures is that the part following DP2 seems to invoke some pitch reset. It should also be noted that although most of the contours have an English equivalent, what sets the hat contour apart from the AB-BA patterns is the sustained pitch plateau between the two pitch accents.



(a) LP-LP pattern

(b) LP-Deaccented pattern

Fig. 5.16: Mean pitch trajectories of active and passive sentences. The black line indicates the mean pitch and the shaded area the standard deviation.



(a) Rise-LP pattern

(b) Rise-EP pattern

Fig. 5.17: Mean pitch trajectories of active and passive sentences.

5.9 Results

A total of 1080 recordings were made (30 participants*36 trials = 1080), 16 recordings had to be excluded due to poor sound quality, leaving us with 1064 good recordings. Eventually, 739 out of 1068 recordings contained no deaccentuations and were used for data analysis.

Figure 5.18 shows the proportions of pitch accents in percentage for each condition. Again, Büring (1997) is mainly interested in cases where both DPs are accented, i.e. without deaccentuation, therefore Figure 5.19 is more informative.

Figure 5.20a gives a visualization of the distribution of the late peak accent across different conditions. Overall, it seems there are more late peaks on the second DP than on the first and the *First DP Focus* condition has the highest proportion of late peak accents. A mixed-effects poisson regression model with random intercepts for speakers was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *Late Peak accents* was taken as independent variable and as dependent variables *focus*, *syntax* and *linear order* were included in the model. The final model performed significantly better than an intercept-only base line model ($\chi^2(5)$: 142.09, p = 0) and is a good, optimal fit (C: 0.9263957, Somers' Dxy: 0.8527914). The model's total explanatory

power is moderate (conditional R2 = 0.23) and the part related to the fixed effects alone (marginal R2) is of 0.13.[12] From the final model it becomes clear that the effect of *linear order* is significant: More late peaks were produced on the second DP than on the first DP (beta = 0.79, 95% CI [0.66, 0.93], p < .001).

The distribution of the early peak accents across the different conditions is shown in Figure 5.20b. The results show us that most early peaks are produced on the second DP and the *First DP focus* condition has the least early peak accents. A poisson regression was run with as independent variable the proportion of *Early Peak accents* and as dependent variables *focus*, *syntax* and *linear order*. No significant results were found for this pitch accent. The final model performed significantly better than an intercept-only base line model (χ^2(5): 128.63, p = 0) and is a good, optimal fit (C: 0.8072192, Somers' Dxy: 0.6114358).[13] The model's explanatory power is substantial (Nagelkerke's R2 = 0.30). From the final model it becomes clear that the effect of *focus* is significant: The fewest Early Peaks have been produced in the *First DP focus* condition.

The distribution of the rising accents across the different conditions is shown in Figure 5.21a. It becomes clear that rising accents were only produced on the first DP. A poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of *Rise accents* was taken as independent variable and as dependent variables *focus*, *syntax* and *linear order* were included in the model. The final model performed significantly better than an intercept-only base line model (χ^2(5): 512.9, p = 0) and is a good, optimal fit (C: 0.8630954, Somers' Dxy: 0.726190). The model's explanatory power is substantial (Nagelkerke's R2 = 0.59).[14] From the final model it becomes clear that *focus* has a significantly effect: Less rising accents were produced in the *First DP focus* condition than in other focus condition (see the full model in Appendix B

[12]The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = First DP focus, is at -0.83 (95% CI [-1.10, -0.57], p < .001).

[13]The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = First DP focus, is at -23.41 (95% CI [-1830.22, 1783.40], p = 0.980).

[14]The model's intercept, corresponding to constituent = First DP, syntax = Active and focus = First DP focus, is at -1.23 (95% CI [-1.68, -0.83], p < .001).

for the estimates). Since there were no productions on the second DP, linear order did not have a significant effect.

Finally, figure 5.21b gives a visualization of the distribution of the early rises across different conditions. Similarly to the rising accents there are no early rises on the second DP. A similar poisson regression to the one for rising accents was run. The final model performed significantly better than an intercept-only base line model ($\chi^2(5)$: 94.283, p = 0) and is a good, optimal fit (C: 0.8217574, Somers ' Dxy: 0.6435149). The model's explanatory power is substantial (Nagelkerke's R2 = 0.32).[15] A significant effect was found for *syntax*: Less early rises were produced in the topic word order than in the active word order (beta = -1.04, 95% CI [-1.91, -0.16], p = 0.020).

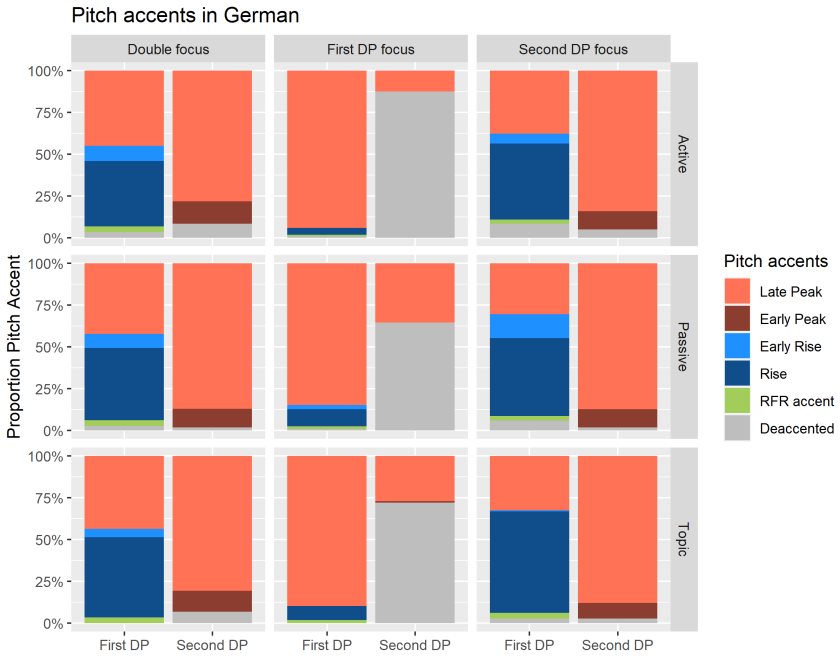


Fig. 5.18: Pitch Accents with Deaccentuation in German

[15]he model's intercept, corresponding to constituent = First DP, syntax = Active and focus = First DP focus, is at -3.38 (95% CI [-4.60, -2.15], p < .001).

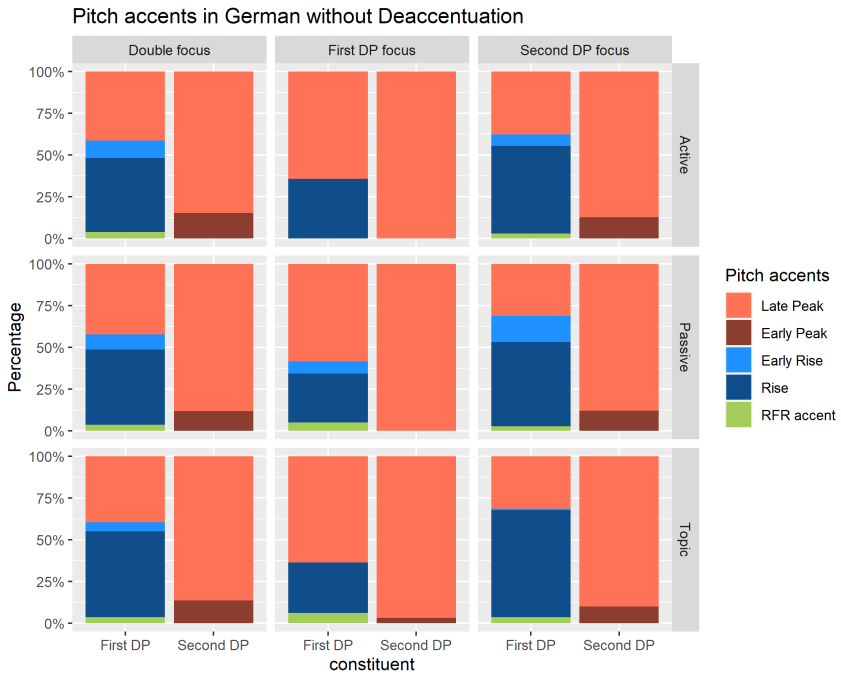
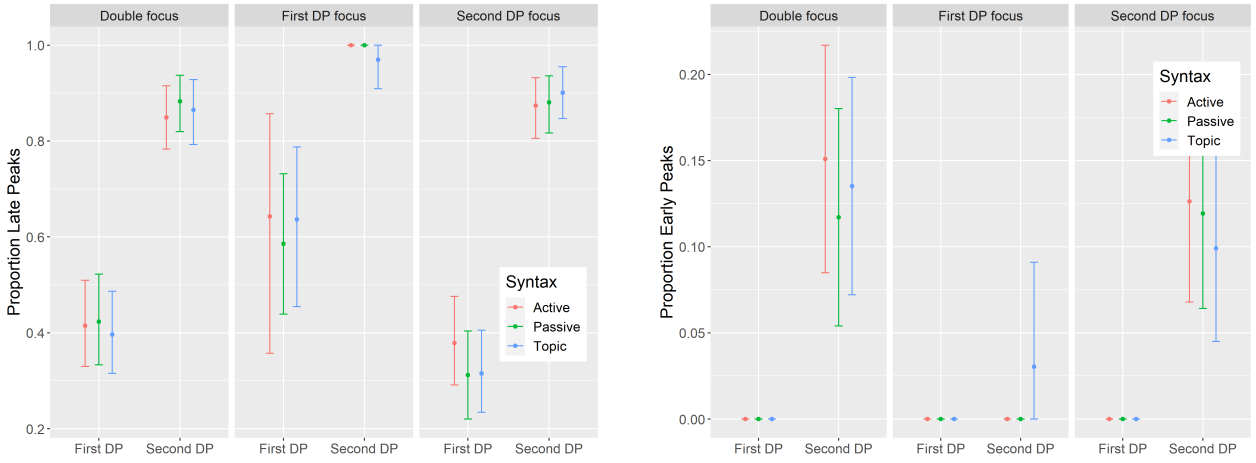


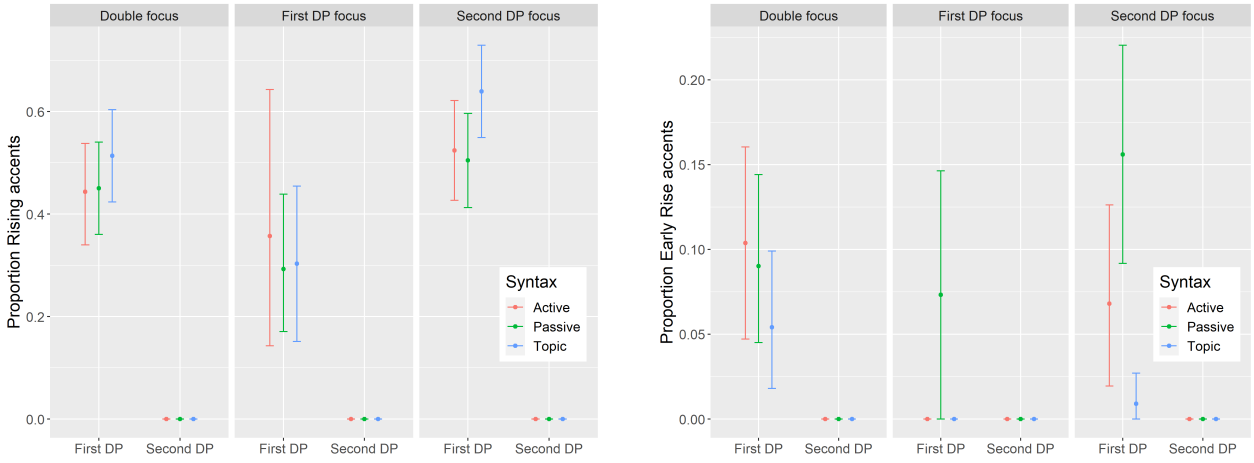
Fig. 5.19: Pitch Accents without Deaccentuation in German



(a) Distribution of Late Peak accents

(b) Distribution of Early Peak accents

Fig. 5.20: Distribution Pitch accents



(a) Distribution of Rising accents

(b) Distribution of Early Rising accents

Fig. 5.21: Distribution Pitch accents

5.10 Conclusion & Discussion

In this chapter, Büring (1997)'s account of the hat contour on sentences with contrastive topics was compared to the Simple-Focus account. The results in this chapter do not support Büring (1997)'s analysis of the hat pattern, i.e. the hat pattern does not convey which constituent is a contrastive topic or a focus.

At first glance one might think that the results do display the expected patterns. For the rising accent, the expectation was that there should be more (early) rising accents on contrastive topics (i.e. more rising accents on the first DP in the *Second DP focus* condition than in the *First DP focus* condition). The results from both the German and Dutch experiment seem to display the expected pattern as can be seen from Figures 5.21a and 5.15a, in addition there was a significant effect of *focus* for the rising accent.[16]

In terms of the falling accent, according to Wagner (2008) (and presumably Büring (1997)) it is only the hat pattern with an early peak that conveys the semantics ascribed to the hat pattern. The expectation was that there should be more early peak accents on focused constituents (i.e. more early peaks on the second DP in the *second DP focus*

[16]It should be noted that the expected pattern was observed for the rising accent, not for the early rising accent that Büring presumably had in mind.

condition than on the second DP in the *First DP focus* condition). A significant effect of *focus* was indeed found for the early peak accent: There are more early peaks on the second DP in the *Second DP focus* condition than in the *First DP focus* condition. The same was also observed for the *Double focus* condition. The results for the *Double focus* condition make sense if one assumes that this condition can be optionally interpreted as containing a contrastive topic. For the *First DP focus* condition no hat patterns are expected and indeed there are almost no early peak accents on the second DP. Interestingly the rising accent and early peak accent seem to be in complementary distribution[17]: Whereas the rising accent does not occur on the second DP, the Early Peak accent does not occur on the first DP. The results for the early peak accent make sense if one assumes that it is only the hat pattern with an early peak that conveys the semantics ascribed to the hat pattern.

However, although the expected pattern was observed for the rising and early peak accents, this pattern seems to be a weak trend rather than a strict correspondence between accent and meaning: Rising accents were indeed more frequently produced on contrastive topics, but a substantial proportion of rising accents was also produced on constituents that would be considered the focus of the sentence. The early peak accent also displayed the expected pattern, but very few actual early peak accents were produced. Moreover, a large proportion of contrastive topics were produced with a falling accent, thus providing even more evidence against the idea that specific pitch accents convey specific meanings.

The Simple-Focus account predicted that the type of accent depended on the linear order in which the accent occurs, with more rising accents on non-final foci. For many pitch accents there was indeed a significant effect of linear order, and rising accents do indeed exclusively occur on the first DP. However, what the Simple-Focus analysis cannot account for is the significant effect of focus contexts for the rising and early peak accents. The Simple-Focus account would have expected no difference across the different focus conditions.

[17]It should be noted that there are much fewer Early Peak accents than rising accents.

In conclusion, none of the different accounts introduced in this chapter can fully account for the production data in this study. Most importantly, there is no evidence to support the claim that specific pitch accents convey specific meanings and thus the results from the current chapter do not support a compositional approach to intonational meaning.

Appendix A: Number of Pitch Accents[18]

	Double focus		First DP focus		Second DP focus	
	First DP	Second DP	First DP	Second DP	First DP	Second DP
Active						
Late Peak	55	92	21	27	65	90
Early Peak	0	9	0	3	0	8
Early Rise	18	0	4	0	20	0
Rise	26	0	4	0	12	0
Rise-Fall-Rise	2	0	1	0	1	0
Passive						
Late Peak	61	99	47	52	60	94
Early Peak	0	5	0	2	0	6
Early Rise	18	0	6	2	19	0
Rise	23	1	4	1	18	0
Rise-Fall-Rise	3	0	0	0	3	0
Topic						
Late Peak	51	88	37	45	51	93
Early Peak	0	13	0	1	0	4
Early Rise	25	0	6	0	26	0
Rise	22	0	1	0	19	0
Rise-Fall-Rise	3	0	2	0	1	0

Table 5.2: Number of contours in sentences without deaccentuation in Dutch

[18] **Go to the OSF-page to access all collected data. Click here.**

	Double focus		First DP focus		Second DP focus	
	First DP	Second DP	First DP	Second DP	First DP	Second DP
Active						
Late Peak	55	92	21	27	65	90
Early Peak	0	9	0	3	0	8
Early Rise	18	0	4	0	20	0
Rise	26	0	4	0	12	0
Rise-Fall-Rise	2	0	1	0	1	0
Passive						
Late Peak	55	92	21	27	65	90
Early Peak	0	9	0	3	0	8
Early Rise	18	0	4	0	20	0
Rise	26	0	4	0	12	0
Rise-Fall-Rise	2	0	1	0	1	0
Topic						
Late Peak	55	92	21	27	65	90
Early Peak	0	9	0	3	0	8
Early Rise	18	0	4	0	20	0
Rise	26	0	4	0	12	0
Rise-Fall-Rise	2	0	1	0	1	0

Table 5.3: Number of contours in sentences without deaccentuation in German

Appendix B: Statistical Models

	Late Peak
(Intercept)	$-0.58\ (0.12)^{***}$
focusFirst DP focus	$0.30\ (0.19)$
focusSecond DP focus	$0.12\ (0.14)$
constituentSecond DP	$0.47\ (0.13)^{***}$
syntaxPassive	$0.04\ (0.15)$
syntaxTopic	$-0.13\ (0.15)$
constituentSecond DP:syntaxPassive	$0.02\ (0.15)$
constituentSecond DP:syntaxTopic	$0.12\ (0.15)$
focusFirst DP focus:syntaxPassive	$0.04\ (0.21)$
focusSecond DP focus:syntaxPassive	$-0.07\ (0.16)$
focusFirst DP focus:syntaxTopic	$0.17\ (0.22)$
focusSecond DP focus:syntaxTopic	$-0.01\ (0.17)$
focusFirst DP focus:constituentSecond DP	$-0.35\ (0.17)^{*}$
focusSecond DP focus:constituentSecond DP	$-0.06\ (0.14)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5.4: Statistical Models Experiment 1.1: Late Peak accents

	Early Peak
(Intercept)	-20.84 (935.85)
constituentSecond DP	18.61 (935.85)
syntaxPassive	-0.53 (0.36)
syntaxTopic	-0.15 (0.33)
focusFirst DP focus	-0.62 (0.45)
focusSecond DP focus	-0.37 (0.30)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5.5: Statistical Models Experiment 1.1: Early Peak accents

	Rise accent
(Intercept)	-2.61 $(0.35)^{***}$
constituentSecond DP	-4.17 $(0.71)^{***}$
syntaxPassive	0.04 (0.21)
syntaxTopic	-0.02 (0.22)
focusDouble focus	1.14 $(0.34)^{***}$
focusSecond DP focus	0.80 $(0.35)^{*}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5.6: Statistical Models Experiment 1.1: Rise accents

	Early Rise
(Intercept)	-2.10 $(0.28)^{***}$
constituentSecond DP	-4.26 $(0.71)^{***}$
syntaxPassive	-0.03 (0.22)
syntaxTopic	0.26 (0.20)
focusDouble focus	0.39 (0.27)
focusSecond DP focus	0.49 (0.27)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5.7: Statistical Models Experiment 1.1: Early Rise accent

	Late Peak
(Intercept)	$-0.88\ (0.15)^{***}$
focusFirst DP focus	$0.44\ (0.37)$
focusSecond DP focus	$-0.09\ (0.22)$
constituentSecond DP	$0.72\ (0.18)^{***}$
syntaxPassive	$0.02\ (0.21)$
syntaxTopic	$-0.05\ (0.21)$
focusFirst DP focus:constituentSecond DP	$-0.27\ (0.47)$
focusSecond DP focus:constituentSecond DP	$0.12\ (0.27)$
focusFirst DP focus:syntaxPassive	$-0.11\ (0.44)$
focusSecond DP focus:syntaxPassive	$-0.21\ (0.31)$
focusFirst DP focus:syntaxTopic	$0.04\ (0.45)$
focusSecond DP focus:syntaxTopic	$-0.14\ (0.32)$
constituentSecond DP:syntaxPassive	$0.02\ (0.26)$
constituentSecond DP:syntaxTopic	$0.06\ (0.26)$
focusFirst DP focus:constituentSecond DP:syntaxPassive	$0.07\ (0.56)$
focusSecond DP focus:constituentSecond DP:syntaxPassive	$0.18\ (0.38)$
focusFirst DP focus:constituentSecond DP:syntaxTopic	$-0.09\ (0.57)$
focusSecond DP focus:constituentSecond DP:syntaxTopic	$0.15\ (0.38)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5.8: Statistical Models Experiment 1.2: Late Peak accents

	Early Peak
(Intercept)	$-20.84\ (935.85)$
constituentSecond DP	$18.61\ (935.85)$
syntaxPassive	$-0.53\ (0.36)$
syntaxTopic	$-0.15\ (0.33)$
focusFirst DP focus	$-0.62\ (0.45)$
focusSecond DP focus	$-0.37\ (0.30)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5.9: Statistical Models Experiment 1.2: Early Peak accents

	Rise accent
(Intercept)	$-1.47\ (0.17)^{***}$
constituentSecond DP	$-4.17\ (0.71)^{***}$
syntaxPassive	$0.04\ (0.21)$
syntaxTopic	$-0.02\ (0.22)$
focusFirst DP focus	$-1.14\ (0.34)^{***}$
focusSecond DP focus	$-0.34\ (0.19)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5.10: Statistical Models Experiment 1.2: Rise accents

	Early Rise
(Intercept)	$-1.71\ (0.18)^{***}$
constituentSecond DP	$-4.26\ (0.71)^{***}$
syntaxPassive	$-0.03\ (0.22)$
syntaxTopic	$0.26\ (0.20)$
focusFirst DP focus	$-0.39\ (0.27)$
focusSecond DP focus	$0.10\ (0.18)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5.11: Statistical Models Experiment 1.2: Early Rise accent

Chapter 6

True Alternatives in German and Dutch

The previous chapter tested Büring (1997)'s claim that the hat pattern is used in sentences containing contrastive topics where a rising accents marks the contrastive topic and a falling accent the focus of the sentence. As we have seen in the previous chapter there is some evidence linking the hat pattern with contrastive topics, although there was no strong evidence that it is the rising accent specifically that is associated with topics: Rising accents were found on foci and falling accents on contrastive topics. The conclusion drawn from the previous chapter is that the hat pattern does not convey which constituent is a contrastive topic or a focus.

The current chapter focuses on the final part of the hat pattern, i.e. the falling accent, and more specifically the semantics that go with two different falling accents, i.e. late peak vs. early peak accents. In this chapter Wagner (2012)'s claims that the hat pattern with an early peak indicates that an alternative proposition must be true, will be put to the test through an online production experiment.

6.1 Early vs. Late Peak

As we have seen from the previous chapter, the contrastive topic manipulations did not seem to fully capture the semantics conveyed by the hat pattern. Wagner (2012) provides

an alternative analysis to the hat pattern which involves making a distinction between hat patterns ending with an early peak and hat patterns ending with a late peak: It is argued that hat patterns with an early peak are most compatible with contexts in which there is a true alternative. Before discussing Wagner (2012)'s account, let us summarize some of the relevant previous literature on the meaning conveyed by early and late peak accents.

6.1.1 Kohler (1991)

Kohler (1991) was one of the first to describe the meaning of the early and late peak contour in German. According to Kohler (1991), the temporal peak alignments of the early and late peak contours form a continuum with the early peak contour (or early peak) on one end of the continuum and the late peak (or late peak) contour on the other end of the continuum. In the middle of this continuum a third category is recognized, i.e. the medial peak contour. Each of these different peak alignments comes with its own particular meaning. Kohler (1991) points out that displacement of the peak to the left induces a semantic change in the 'new vs. established' dimension, whereas displacement to the right produces a change in the degree of emphasis. Perception experiments show that a displacement of the peak to the left induces a categorical change from the early to the middle peak, whereas only a gradual auditory change is perceived from a middle to a late peak Féry (2010), the same conclusion was drawn by Rathcke and Harrington (2006).

Based on a survey in which participants where asked to describe the meaning of certain contours, Kohler (1991) abstracted and paraphrased the meanings of each contour. Generally speaking, the early peak contour is used for established facts and it conveys that there is no room for discussion, it is often used as a final summing up of an argument. The medial peak contour is used for new facts and it conveys that there is room for discussion, it is often used when starting a new argument. Finally, the late peak puts emphasis on a new fact and contrast it to what is part of both the speaker's or hearer's common ground, i.e. set of shared beliefs that are taken to be true by both listener and

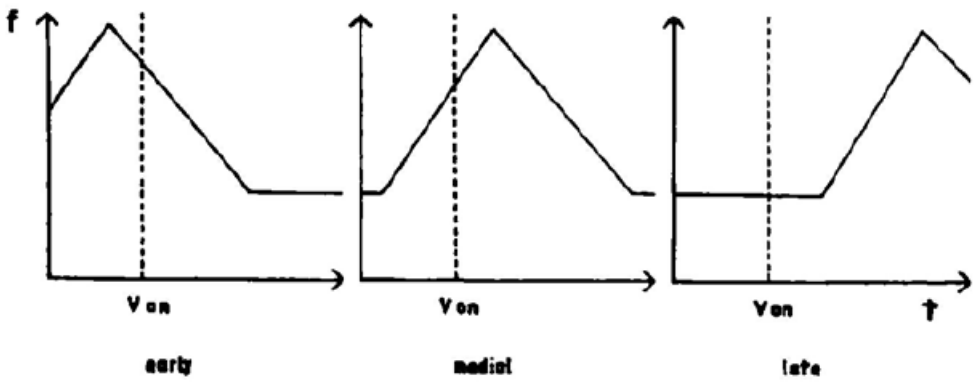


Fig. 6.1: Early, medial and late peak according to the Kiel intonation model (from (Gartenberg and Panzlaff-Reuter 1991)). The dashed line Von′ marks vowel onset

speaker. Based on Rathcke and Harrington (2006) and Féry (2010) who only acknowledge two peak accents in German, this study also only acknowledges two peak accents: The early and late peak accent (with the Kohler (1991) medial peak corresponding to a less emphasized version of the late peak accent).

Although these are believed to be the basic meanings of each contour, the actual meanings associated with individual utterances depend on the interplay of these basic semantics of intonation contours with the semantics of different syntactic structures and of the lexicon (Kohler 1991).

If an early peak is used in questions, then the question receives special connotations in keeping with the semantics of the early peak contour: the question is asked with the speaker already having presumed knowledge of the answer. An example would be a teacher asking “Who did it?", the early peak contour in this case would convey a possible threat, i.e. “I will figure out whoever did it". If an early peak contour is used with imperatives, there is again a contradiction between the signalling of the expected completion of an action and the order to carry it out, signalled by the syntax. This contradiction produces the connotation of annoyance and impatience, e.g. “Mach bitte das FENSTER zu." (“Shut the window, please.") with an early peak contour the utterance sounds rather annoyed compared to a medial peak contour.

In alternative questions an early peak in second position signals a choice within a closed set of alternatives, whereas a succession of medial peak with low F0 in between refers to an open set of alternatives, which are simply given as possible examples from a longer list. For instance “Willst du Tee oder Kaffee?" ("Would you like tea or coffee?") with an early peak contour limits the set of alternatives just to *tea* or *coffee*, whereas with a medial peak contour the alternative set would consist of *tea* or *coffee* or any other alternative to these.[1]

Kohler (1991) concludes by claiming that there is a direct link between particular f0 contours and specific meanings, these specific meanings however interact with various other levels of meaning (i.e. syntax, lexicon etc.) thus resulting in an apparent variety of meanings for one particular f0 contour.

6.1.2 Keijsper et al. (1984)

Just like Kohler (1991) was one of the first to extensively describe the meaning of the early peak contour in German, Keijsper et al. (1984) was the first to do so for Dutch even before Kohler (1991). It will become clear that the early peak contours in the two languages have very similar meanings.

Unlike Kohler (1991), Keijsper et al. (1984) recognizes only two falling f0 contours in Dutch, namely a falling f0 contour and a so-called pointed hat. Although there is no detailed phonetic description of the two contours, but merely a rather unclear schematic representation, a more detailed description is given by Caspers (1999). According to Caspers (1999), the former contour would correspond to an early peak contour with a high initial boundary tone (%H H!*+L) and the latter to a late peak contour with a low initial boundary tone (%L H*+L): I will henceforth assume that these are the intended phonological categories Keijsper et al. (1984) refers to.

[1] Furthermore Kohler (1991) argues that a rising contour, like a medial peak contour, conveys the same open alternative set but it sounds less authoritative and more friendly.

The early peak contour is used with established information, or “projected information" as Keijsper et al. (1984) calls it: With the term “projected" Keijsper et al. (1984) means that the existence of the referent has been projected before the moment of speaking, this concept is basically the same as Kohler (1991)’s idea of an established fact. More specifically, the early peak indicates that the existence of a certain proposition was known before the moment of speaking i.e. it was part of the common ground between two speakers. Let us see more clearly what this means exactly.

(61) *Dutch* (Keijsper et al. 1984: 28)

a. Een stropdas
%H H!*+L
‘A bow tie!’

b. Mijn huis staat in brand
%H H!*+L
‘My house is on fire.’

Suppose it is your birthday and you received a birthday gift and the moment you open your present you say “A bow tie" with an early peak contour (61a): This would be seen as rather impolite since you convey that you already knew you would receive a bow tie even before you opened the present. It sounds as if you are disappointed since you expect to receive a surprise, i.e. something unexpected, rather than something you anticipated.

A slightly different situation is illustrated in Example (61b). Saying this sentence with an (early) peak contour somehow conveys that you have always known that at some point your house would catch on fire, you anticipated the moment this would happen and are not surprised about it. These two examples should have made clear what is meant with *projection before the moment of speaking*: It means that the information conveyed did not come as a surprise and was somehow expected. This is a rather vague notion and not semantically precise, as we will see later on there are more precise descriptions of the (early peak) hat pattern that supposedly better capture its semantics.

Again, when used on the last element in an alternative question the early peak refers to a closed set of alternatives consisting of only the referents introduced in the utterance. For example in “Wil je koffie of thee?" (“Do you want coffee or tea?") the alternative set

consists of only *coffee* and *tea* if *thee* is produced with an early peak contour, this is in contrast to when *thee* is produced with a rising contour in which case there is an open set of alternatives and the options for the listener to choose from is not just limited to *coffee* or *tea*.

The late peak contour is used when there is new information and it can be used to place emphasis on an element. Moreover it is the most neutral means of assigning prominence on a syllable, in other words the late peak contour is the most neutral option for accented syllables. Only in final position a further meaning aspect is added, namely it marks focused information as new. As should have become clear the two contours in German and Dutch have very similar meanings.

6.1.3 Caspers et al. (1998)

Caspers et al. (1998) tested several factors to specifically distinguish the early peak contour with high initial boundary tone (%H H!*+L) from the late peak contour with low initial boundary tone (or the 'pointed hat' contour, %L H*L). According to Gussenhoven (1991) both contours mark information as new, but the early peak contour sounds more irritated: Assuming that it is not unusual for sentences containing predictable information (or "projected" information in Keijsper et al. (1984)'s terminology) to be uttered with an irritated 'tone', this analysis would provide an alternative explanation for the finding that the early peak contour is acceptable only on 'predictable' sentences (Caspers et al. 1998). Caspers et al. (1998) attempts to tease apart the two notions of "projection/expectation" and "irritation", to see which notion better accounts for the meaning difference between the two contours or whether both notions are needed independently. Different contexts were created to capture the notion of "projection".

The target sentence forms either the end of an enumeration (62) —i.e., no speaker-turn has occurred and the focused information is known by the same speaker before the moment of speaking —or it is the answer to a question (63) —i.e. a speaker-turn has

occurred immediately before. The two contexts are produced with both contours and presented to the participants. Participants were asked to judge the acceptability of each context given the contour as well as how irritated it sounded.

(62) **Enumeration context** (Caspers et al. 1998: p. 2)

Together with a colleague you are visiting Amsterdam with a class. You have divided jobs and are meeting again for lunch. To the question of your colleague how things are going you answer: "Not too well, because the exhibition was disappointing, Jolanda was robbed, and..." : "Marina is missing" .

(63) **Question-Answer context** (Caspers et al. 1998: p. 2)

You are visiting the Rijksmuseum with a number of pupils and a colleague. When leaving the museum you notice that a pupil is missing. You are busy finding out if other pupils know where she is as your colleague comes outside. On his question what is going on you answer: "Marina is missing" .

The results clearly show that the early peak contour is not appropriate as an answer to a question, whereas both contours can be used to mark the last element in an enumeration, but with a light preference for the early peak contour (6.2). Furthermore, the early peak contours sounds more irritated than the late peak contour with low onset.

Caspers et al. (1998) concludes that both notions of "projection" as defined by Keijsper et al. (1984) and "irritation" play a role in explaining the semantic difference between the early and late peak contour. However, there is a confound in the choice of contours used in this study: It is unclear whether the meaning difference is due to a difference in onset, i.e. a high onset (or initial boundary tone) for the early peak contour and a low onset for the late peak contour, or a difference in the actual pitch accent, i.e. an early peak vs. a late peak contour.

Grabe et al. (1998) show that a high onset followed by a high pitch accent (e.g. %H !H*L) sounds less favorable —less friendly, less polite, more irritated, and more aloof —

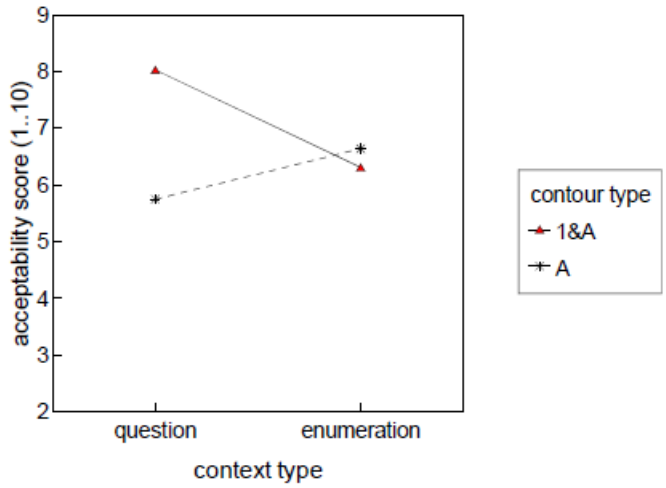


Fig. 6.2: Acceptability scores per contour type, broken down by context type.

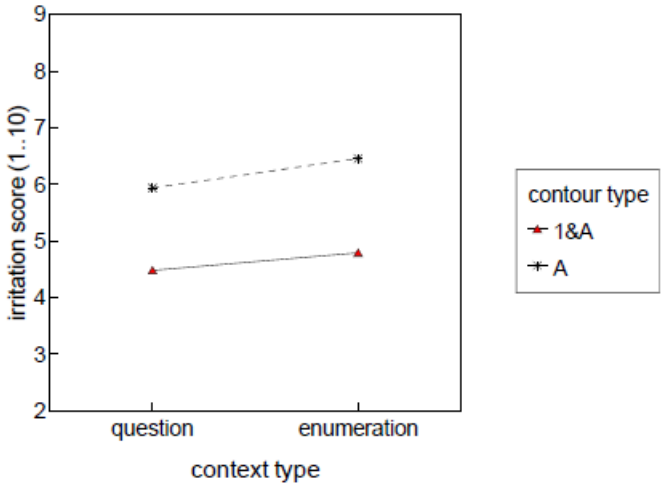


Fig. 6.3: Irritation scores per contour type, broken down by context type.

than a low onset followed by a high pitch accent (e.g. %L H*L). It is therefore very likely that the higher levels of irritation are due to the high onset rather than the early peak pitch accent itself.

To get around this confound Caspers (1999) set up a similar experiment testing for several more factors to distinguish the early peak (6.4) from a late peak with a high initial boundary tone (6.5) and a pointed hat pattern, i.e. late peak with low initial boundary tone (6.6). She looked into the following parameters: New vs. projected (i.e. previously referred to) information, levels of irritation, detachment and finality.

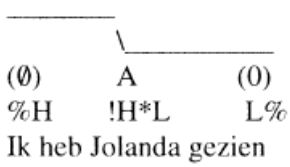


Fig. 6.4: early peak contour with low onset/ initial boundary tone with transcriptions in ToDI and GDI (Grammar of Dutch Intonation 't Hart et al. (1990))

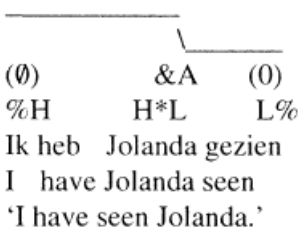


Fig. 6.5: late peak contour with high onset/ initial boundary tone with transcriptions in ToDI and GDI (Grammar of Dutch Intonation 't Hart et al. (1990))

It was hypothesized that a difference in the timing of the fall corresponds to a difference between new and projected information (i.e. previously referred to information or anticipated/expected information). More specifically the early peak is expected to be compatible with projected information, while the late peak is associated with new information.

In a rating experiment, subjects were asked to assess different contexts and contours on a ten-point scale. The materials were presented in the form of a series of short conversations between two teachers. One such example conversation is illustrated below (64). Answer B1 (64a) contains all new information, therefore the late peak is expected to fit this context better than the early peak. Answer B2 (64b) is the one in which the focused sentence has been projected before the moment of speaking. In answer B2 there are two

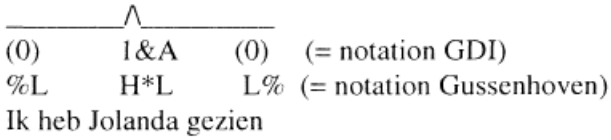


Fig. 6.6: late peak contour with low onset/ initial boundary tone or "pointed hat"

events that are related to each other, both are namely actions undertaken by *Jolanda* to approach someone for the cleaning squad: One event is that she asked Jan-Willem and the other event is she called Marina, both are things Jolanda has done in order to approach someone for the cleaning squad. In other words one related event (i.e. Jolanda asking Jan-Willem) has been projected before the other event (i.e. Jolanda calling Marina). Caspers (1999) predicts that the late peak should be more acceptable with an answer like B1 than the early peak.

(64) *Dutch* (Caspers 1999: 31)

A: Jolanda heeft nog niemand benaderd voor de opruimploeg.
A: Jolanda has yet nobody approached for the cleaning.squad

'Jolanda hasn't approached anyone for the cleaning squad yet.'

a. B1: Jawel! *Ze heeft Marina gebeld*.
B1: Yes She has Marina called.
'Yes she has! She has called Marina'

b. B2: Jawel! Ze heeft Jan-Willem gevraagd, en *ze heeft Marina gebeld*.
B2: Yes she has Jan-Willem asked, and she has Marina called.
'Yes she has! She has asked Jan-willem and she has called Marina'

From the results, it was concluded that the early peak sounded more detached, more irritated, more final and less acceptable than the late peak in general. One recurring result that has been found and mentioned in previous studies is that the early peak contour sounds more irritated than the late peak or "pointed hat" contour. Moreover, Caspers (1999) found that the early peak did not go well with new information compared both late peak contours. This can be clearly seen from Figure (6.7): In the *new* + condition[2], the pointed hat pattern (1&A) is the most acceptable contour, but more importantly the late peak (&A) is clearly more acceptable than the early peak (A). In the *projected* condition there is no clear difference between the early and late fall.[3]

[2]The plus and minus signs indicate whether there was elipsis or not, but this has no importance for the current study

[3]For the purpose of the current study we are not discussing the *new* - condition, i.e. new information containing elipsis

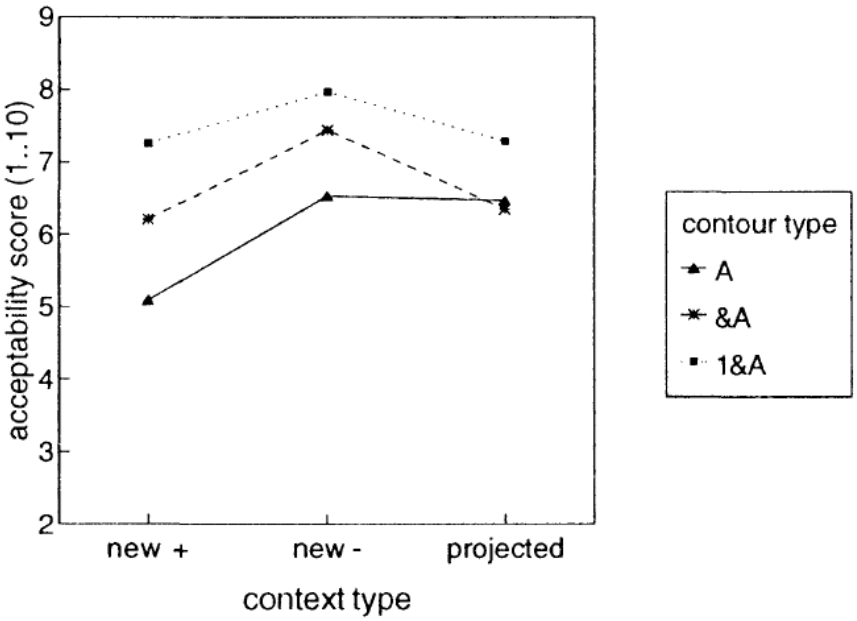


Fig. 6.7: Results Caspers (1999).

Caspers (1999) argues that these findings support Keijsper et al. (1984)'s proposal and her hypothesis according to which the early peak indicates that the existence of the focused information has been projected (i.e. expected) before the moment of speaking.

However, it does not follow from Casper's results that the early peak is associated with projected information since both early (A) and late falls (&A) are equally acceptable in this condition. These results therefore do not directly support Keijsper's analysis, the only conclusion that can be drawn from the results is that the early peak does not go well with new information: This finding is not new in the sense that similar results were found in her previous experiments (Caspers 1997; Caspers et al. 1998; Caspers 2000).

Something else to question is whether Casper's stimuli are testing for a new versus expected information status or whether something else is being tested. Keijsper et al. (1984)'s definition of projected information says that projected information is information that has been expected/anticipated previous to the moment of speaking. It is not particularly clear how the act of *Jolanda calling Marina* is more expected in Answer B2 than it is in answer B1. The fact that she has asked *Jan-Willem*, does not make it more likely that she would

have called Marina. The relation between these two proposition rather seems to be that both are part of a set of alternatives. In other words, both are part of the set of potential solutions to finding someone for the cleaning squad. The idea that there is a connection between the early peak and a set of alternatives will be tested in this chapter.

Formulating the meaning of the early peak in terms of alternative propositions has the advantage that it can be expressed more formally, this is in contrast to the notion of projected information which is still rather vague as there is no formal way of representing it.

6.2 Topic/Focus Analysis

As extensively discussed in the previous chapter Büring (1997) takes the hat pattern to be compositional with the initial rise marking a contrastive topic and the final fall marking a focus. Formally whichever element in the sentence receives focus will be part of the set of propositions that are considered well-formed alternatives, this set is referred to as the focus value. Like focus, contrastive topics also induce alternatives and are represented as a set containing different focus values, i.e. a set of sets of propositions, this is referred to as the *topic value* of a sentence. It should become clear that the focus value is nested within the topic value, by changing the subject, the topic of the sentence changes. Crucial to Büring (1997)'s analysis of the hat contour is that contrastive topics come with a so-called disputability implicature. This implicature is suggests that in sentences with a contrastive topic there is a question in the set of questions denoted by the topic value which is still considered disputable.

Büring (1997)'s analysis of the hat pattern was however not supported by the results from the previous chapter: The hat pattern does not convey which constituent is a contrastive topic or a focus.

6.3 True Alternative Propositions

Wagner (2008; 2012) provides an alternative analysis to the hat pattern in German which involves making a distinction between early and late peak hat patterns. As we have seen from the previous sections Kohler (1991) argues that the early peak accent in German conveys that something is an 'established' fact whereas the late peak corresponds to new facts. This corresponds to Keijsper et al. (1984)'s analysis for the early peak accent in Dutch which according to Keijsper et al. (1984) conveys that some proposition is 'expected' in the discourse. Like Kohler (1990), Keijsper et al. (1984) takes the late peak accent to be most compatible in contexts establishing new information.

The results from the experiments run by Caspers (1997); Caspers et al. (1998); Caspers (2000) provided evidence showing that early peak accents are less compatible with new information. In terms of the link between early peaks and expected infomration however, I argued that there was insufficient convincing evidence showing that early peak accents convey that some proposition is expected in the discourse. The notions of 'established' fact and expected information are rather vague and hard to formalize, a more formal approach to the meaning of early and late peak accents would be more desirable and easier to implement into an experiment. Wagner (2012) provides such a more formal analysis of early peak hat patterns. Although Wagner (2012)'s analysis applies to the hat pattern rather than the individual early/late peak accents, Wagner argues that the meaningful component of the hat pattern lies in the falling gesture rather than the rising gesture. In fact, the rising gesture is not obligatory and might be completely absent in cases of phonetic reduction. Therefore the meanings ascribed to the early peak hat contour should arguably overlap with (if not equate to) the meanings ascribed to the early peak accent.

In contrast to Büring (1997), Wagner (2012) and along with him many other researchers argue for a multiple focus analysis to the hat pattern rather than a topic/focus analysis (Van Hoof 2003; Ludwig 2006; Wagner 2008; 2012). For Ludwig (2006), the German hat pattern can be interpreted as an operator which takes two foci as its argument and means

something like "*there is at least one true proposition (sentence) that is the result of replacing both foci with respective alternatives*". Elaborating on Ludwig (2006), Wagner (2008; 2012) argues that the hat-contour presupposes that an alternative proposition is true. More specifically, it is only the hat pattern with an early peak as opposed to a late peak accent, that conveys that a contextually alternative proposition is true besides the proposition expressed bearing the hat pattern. It follows from Wagner's definition of the hat pattern that it should be less acceptable when there are no alternative propositions or when it is unclear if there are alternative propositions. The hypothesis that will be tested in this chapter is Wagner (2012)'s claim that the early peak hat pattern conveys that a contextually alternative proposition is true besides the proposition expressed bearing the hat pattern.

6.4 Methodology

In order to test the claims made by Wagner (2012), an online production experiment was set up using jsPsych (de Leeuw 2015). Participants were presented with a number of carefully manipulated questions, both in written form as well as auditorily, after which they were asked to record their answer and assess how well the answer sounded.

6.4.1 Participants

A total of 60 participants were recruited through Prolific, 30 of them were native speakers of Standard Dutch (as spoken in the Netherlands) and the other 30 were native speakers of Standard German (as spoken in Germany). A compensation of 2.33 C$ was given which boils down to an hourly rate of 13.98 C$ since the experiment itself took about 10 minutes.

6.4.2 Stimuli

Each stimulus consists of a short context, a question and an answer. The short contexts determine the number of possible alternatives. The questions consist of simple polar ques-

tions. For the answers there are five different conditions (see Figure 65): Answers with an *explicit true alternative* as illustrated in (65a), in this case *Valentina should call Franco* is a true alternative to *John should call Emma* and the affirmative *yes* ensures that the alternative is explicit. In the *no alternative correction* condition, the negative particle *no* excludes *Valentina should call Franco* as an alternative to *John should call Emma*, the latter instead is a correction to the former (65e). In the *implicit true alternative* condition, there is a true alternative to *John should call Emma* as indicated by the short context *Two phone calls should be made*, but it is not clear what the alternative is as indicated by the phrase *I don't know who should make the second phone call* (65b). The *open alternative* conditions makes no claims on whether there are any alternatives or not (65c). Finally, in the *no alternative question* condition, the participant is presented with a mutually exclusive choice between two alternative, therefore only one can be true and there is no alternative proposition true (65d)

(65) a. (Explicit True Alternative)

A: Twee telefoontjes zouden moeten worden gepleegd. Zou Valentina Franco moeten opbellen?

B: Ja. En ik weet nog iets: Johan zou Emma moeten opbellen.

"A: Two phone calls should be made. Should Valentina call Franco?"

"B: Yes. And I know what else: John should call Emma."

b. (Implicit True Alternative)

A: Twee telefoontjes zouden moeten worden gepleegd. Zou Valentina Franco moeten opbellen?

B: Geen idée. Maar een ding is zeker: Johan zou Emma moeten opbellen. Ik weet niet wie het tweede telefoontje zou moeten plegen.

"A: Two phone calls should be made. Should Valentina call Franco?"

B: I don't know. But one thing is clear: John should call Emma. I don't know who should make the second phone call.

c. (Open Alternative)

A: Ik vraag me af of nog enkele telefoontjes moeten worden gepleegd. Zou Valentina Franco moeten opbellen?

B: Geen idée. Maar een ding is zeker: Johan zou Emma moeten opbellen. Misschien is dat genoeg?

A: I wonder if any phone calls should be made. Should Valentina call Franco?

B: I don't know. But one thing is clear: John should call Emma. Maybe that'll be good enough?

d. (No Alternative Question)

A: Zou Johan Emma moeten opbellen of zou Emma Johan moeten opbellen?

B: Nou dat is wel duidelijk: Johan zou Emma moeten opbellen. Dat moet genoeg zijn.

A: Should John call Emma or should Emma call John?

B: Well, that's pretty clear: John should call Emma. That should be more than enough.

e. (No Alternative Correction)

A: Een telefoontje zou moeten worden gepleegd. Zou Valentina Franco moeten opbellen?

B: Nee. Johan zou Emma moeten opbellen. Dat moet genoeg zijn.

A: One phone call should be made. Should Valentina call Franco?

B: No. John should call Emma. That should be more than enough.

6.4.3 Procedure

The stimuli were presented to the participant in a latin-square design. Participants are directed to the experiment via an online link. The main experiment starts off with a fixation point, after which the context and question are given in written form and finally the context and question are presented auditorily. Each question ends with a beep after which the participant is asked to record the question as natural as possible. The trial finishes off with an acceptability input screen in which the participant has to indicate how natural their answer sounded, based on a scale from 1 to 8 where 1 means completely unnatural and 8 means completely natural.

6.4.4 Expectations

Büring (1997) expects most hat patterns, i.e. (early) rising accents followed by an early peak accent, to occur in contexts with open alternatives, this follows from the fact that hat patterns come with a disputability implicature. Wagner (2012) on the other hands expects hat patterns with early peaks to be most compatible in contexts with true alternative propositions, be it explicit or implicit. For all other conditions both Büring (1997) and Wagner (2012) would mostly expect LP-LP contours given that there is a double contrast in each condition. An overview of the different accounts is provided in Table 6.1.

	Explicit True	**Implicit True**	**Open Alter**	**Alter Question**	**Correction**
Büring	LP-LP	LP-LP	EP Hat	LP-LP	LP-LP
Wagner	EP Hat	EP Hat	LP-LP	LP-LP	LP-LP

Table 6.1: Expectations hat pattern different accounts

6.5 Results

6.5.1 Dutch

A total of 900 recordings were made (30 participants*30 trials = 1080), 40 recordings had to be excluded due to poor sound quality, leaving us with 860 good recordings. Eventually, 734 out of 860 recordings contained no deaccentuations and were used for data analysis. For the results the recorded sentences were annotated by the author in terms of the pitch accents on the first DP and second DP. The following categories were distinguished for the annotation: late peak, early peak, Rise-Fall-Rise, early rise, rise, Deaccented. For a phonological description of each pitch accent see Chapter 5.

Figure 6.8 shows the proportions of pitch accents in percentage for each condition. Once again we are mainly interested in cases where both DPs are focused, i.e. without deaccentuation, therefore Figure 6.9 where all cases with deaccentuation have been excluded, is more informative: The exact numbers can be found in the Appendix.

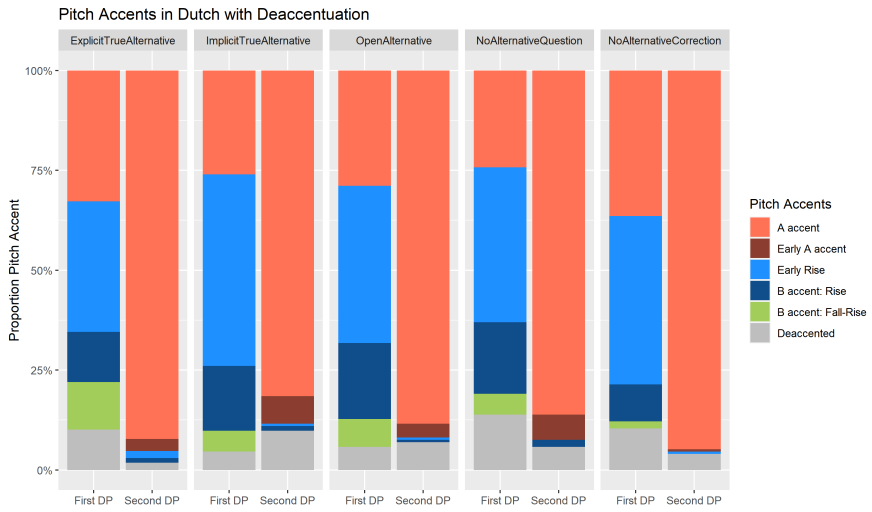


Fig. 6.8: Pitch Accents with Deaccentuation in Dutch

Figure 6.10 gives a visualization of the distribution of the late peaks across different conditions. From the figure it becomes clear that more late peaks have been produced on the second DP than on the first DP. A poisson regression model was fitted to the data

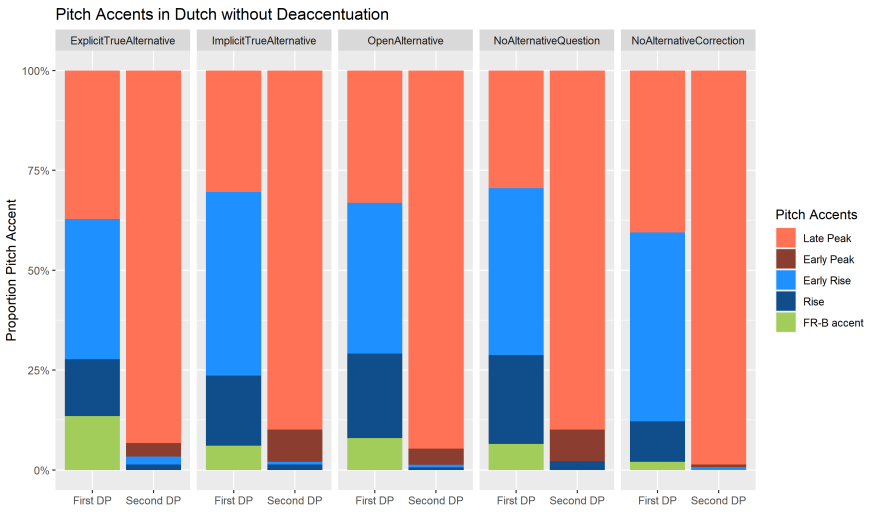


Fig. 6.9: Pitch Accents without Deaccentuation in Dutch

in a step-wise-step up procedure. In the final minimal adequate model the proportion of late peak accents was taken as independent variable and as dependent variables the *alternative* condition and *linear order* condition. The final model performed significantly better than an intercept-only base line model (χ^2(5): 270.11, p = 0) and is a good, optimal fit (C: 0.8306038, Somers' Dxy: 0.6612076). The model's explanatory power is substantial (Nagelkerke's R2 = 0.31).[4] From the model comparison it becomes clear that there is a significant effect of *linear order*: More late peak accents were produced on the second DP than on the first DP (beta = 1.09, 95% CI [0.95, 1.24], p < .001).

Figure 6.11 gives a visualization of the distribution of the early peak accents. This Figure shows us that no early peaks are produced on the first DP and less early peak seem to have been produced in the *NoAlternativeCorrection* condition. A similar poisson regression to the late peak one was performed. The final model performed significantly better than an intercept-only base line model (χ^2(5): 62.181, p = 0) and is a good, moderate fit (C: 0.8361424, Somers' Dxy: 0.6722849). The model's explanatory power is moderate (Nagelkerke's R2 = 0.24). No significant results were reported by the model.

[4]The model's intercept, corresponding to constituent = First DP and assertion = ExplicitTrueAlternative, is at -1.16 (95% CI [-1.34, -0.99], p < .001).

Figure 6.12 gives a visualization of the distribution of the rise accent. No clear pattern can be observed from the figure except that more rise accents have been produced on the first DP than on the second DP. Again, a poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of rise accents was taken as independent variable and as dependent variables the *alternative* condition and *linear order* condition. The final model performed significantly better than an intercept-only base line model (χ^2(5): 138.24, p = 0) and is a good, moderate fit (C: 0.7764593, Somers' Dxy: 0.5529186). The model's explanatory power is moderate (Nagelkerke's R2 = 0.23).[5] From the model comparison it becomes clear that there is a significant effect of *linear order*: Less rising accents were produced on the second DP than on the first DP (beta = -2.78, 95% CI [-3.58, -2.13], p < .001).

The distribution of the early rises is shown in Figure 6.13. Again more early rises are observed on the first DP than on the second DP and it seems that less early rises are produced in the *ExplicitTrueAlternative* condition than in the other conditions. A similar poisson regression to the rising accent one was performed and similar results were obtained. The final model performed significantly better than an intercept-only base line model (χ^2(5): 431.71, p = 0) and is a good, optimal fit (C: 0.8232822, Somers' Dxy: 0.6465643). The model's explanatory power is substantial (Nagelkerke's R2 = 0.46).[6] From the model comparison it becomes clear that there is a significant effect of *linear order*: Less early rising accents were produced on the second DP than on the first DP (beta = -4.05, 95% CI [-4.98, -3.34], p < .001).

Finally, figure 6.14 gives a visualization of the distribution of the Rise-Fall-Rise accents. Interestingly, like the English data, in Dutch most RFR accents are produced in the *ExplicitTrueAlternative* condition. A similar poisson regression to previous ones was performed. The final model performed significantly better than an intercept-only base

[5]The model's intercept, corresponding to constituent = First DP and assertion = ExplicitTrueAlternative, is at -2.05 (95% CI [-2.49, -1.66], p < .001).

[6]The model's intercept, corresponding to constituent = First DP and assertion = ExplicitTrueAlternative, is at -1.08 (95% CI [-1.35, -0.83], p < .001)).

line model (χ^2(5): 88.94, p = 0) and is a good, optimal fit (C: 0.8313318, Somers' Dxy: 0.6626637). The model's explanatory power is substantial (Nagelkerke's R2 = 0.26). The model reports that only the *alternative* manipulation has a significant effect: There are significantly more RFR accents in the *ExplicitTrueAlternative* condition than in the other *alternative* conditions, except for the *OpenAlternative* condition which was not significantly different from the baseline (see the full model in the appendix for the actual estimates)

Fig. 6.10: Distribution of late peak accents in Dutch

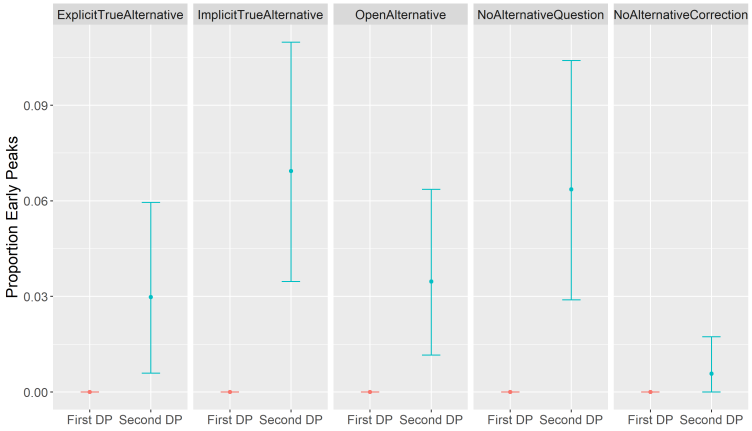


Fig. 6.11: Distribution of early peak accents in Dutch

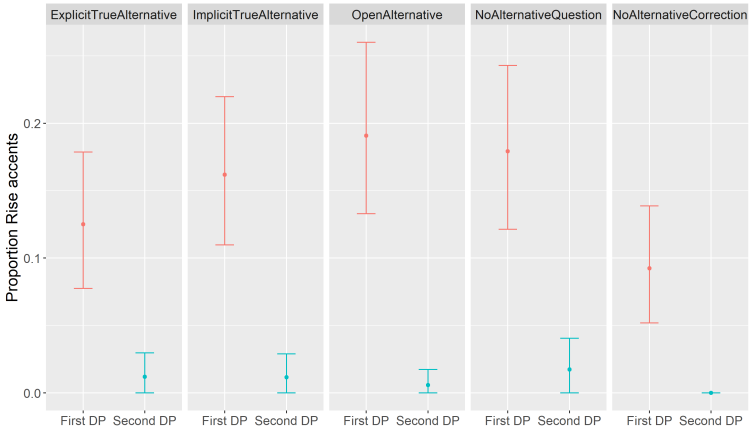


Fig. 6.12: Distribution of Rising accents in Dutch

Fig. 6.13: Distribution of Early Rising accents in Dutch

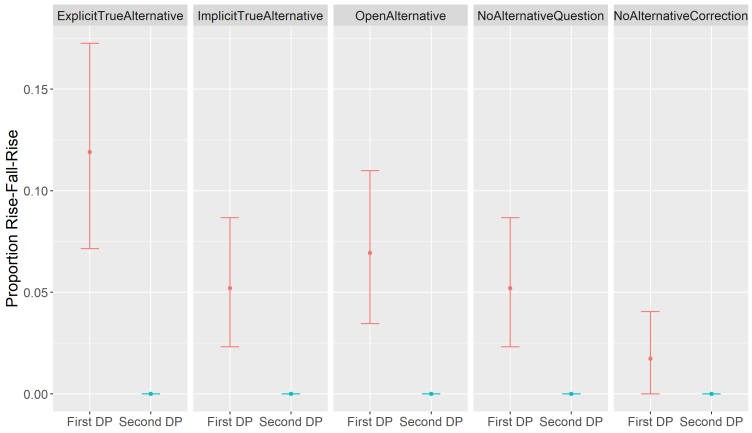


Fig. 6.14: Distribution of Rise-Fall-Rise B accents in Dutch

6.5.2 German

A total of 900 recordings were made (30 participants*30 trials = 1080), two participants had to be excluded since no recordings were made and an additional 14 recordings had to be excluded due to poor sound quality, leaving us with 816 good recordings. Eventually, 711 out of 816 recordings contained no deaccentuations and were used for data analysis.

Figure 6.15 shows the proportions of pitch accents in percentage for each condition. Once again we are mainly interested in cases where both DPs are focused, i.e. without deaccentuation, therefore Figure 6.16 where all cases with deaccentuation have been excluded, is more informative.

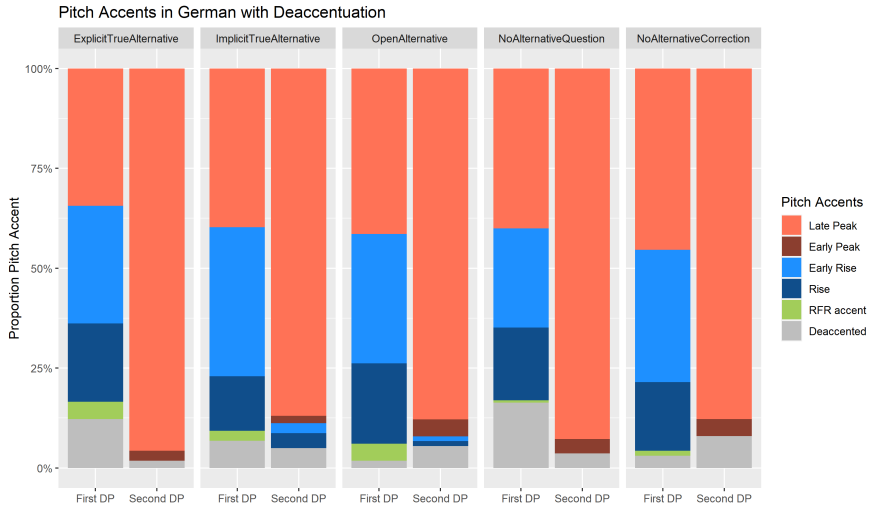


Fig. 6.15: Pitch Accents with Deaccentuation in German

Figure 6.17 gives a visualization of the distribution of the late peak accent. From this figure it becomes clear that more late peaks have been produced on the second DP than on the first DP. A poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of late peak accents was taken as independent variable and as dependent variables the *alternative* condition and *linear order* condition. The final model performed significantly better than an intercept-only base line model ($\chi^2(5)$: 160.88, p = 0) and is a good, not optimal fit (C: 0.7826068, Somers' Dxy: 0.5652137). The model's explanatory power is moderate (Nagelkerke's R2

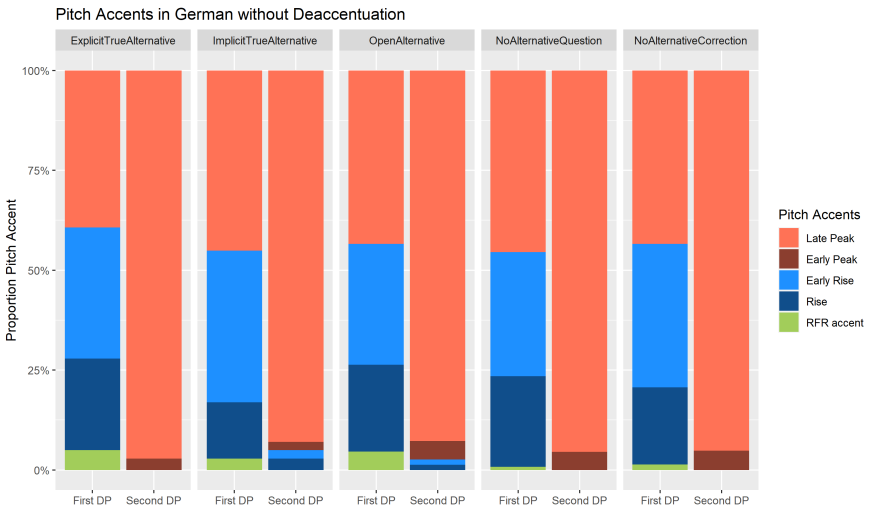


Fig. 6.16: Pitch Accents without Deaccentuation in German

= 0.22).[7] From the model comparison it becomes clear that there is a significant effect of *linear order*: More late peak accents were produced on the second DP than on the first DP (beta = 0.81, 95% CI [0.68, 0.94], p < .001).

Figure 6.18 gives a visualization of the distribution of the early peak accents. Other than no early peaks on the first DP, no clear pattern arises from the production data. A similar poisson regression to the late peak one was performed. The final model performed significantly better than an intercept-only base line model (χ^2(5): 39.941, p = 0) and is a good, moderate fit (C: 0.7951194, Somers' Dxy: 0.5902388). The model's explanatory power is moderate (Nagelkerke's R2 = 0.19). No signficiant results were reported by the model.

Figure 6.19 gives a visualization of the distribution of the rise accent. From this figure it becomes clear that rising accents are mainly produced on the first DP. Again, a poisson regression model was fitted to the data in a step-wise-step up procedure. In the final minimal adequate model the proportion of rise accents was taken as independent variable and as dependent variables the *alternative* condition and *linear order* condition. The final model performed significantly better than an intercept-only base line model (χ^2(5): 150.35,

[7]The model's intercept, corresponding to constituent = First DP and assertion = ExplicitTrueAlternative, is at -0.91 (95% CI [-1.08, -0.75], p < .001).

p = 0) and is a good, not optimal fit (C: 0.759462, Somers' Dxy: 0.5189251). The model's explanatory power is moderate (Nagelkerke's R2 = 0.25).[8] From the model comparison it becomes clear that there is a significant effect of *linear order*: Less rising accents were produced on the second DP than on the first DP (beta = -2.90, 95% CI [-3.70, -2.25], p < .001).

The distribution of the early rises is shown in Figure 6.20. Like the rising accents, early rises are mainly produced on the first DP. A similar poisson regression to the rising accent one was performed and similar results were obtained. The final model performed significantly better than an intercept-only base line model (χ^2(5): 312.21, p = 0) and is a good, optimal fit (C: 0.8092954, Somers' Dxy: 0.6185909). The model's explanatory power is substantial (Nagelkerke's R2 = 0.39).[9] From the model comparison it becomes clear that there is a significant effect of *linear order*: Less early rising accents were produced on the second DP than on the first DP (beta = -3.75, 95% CI [-4.68, -3.04], p < .001).

Finally, figure 6.21 gives a visualization of the distribution of the RFR accents. This figure shows us that no RFR accents are produced on the second DP. A similar poisson regression to previous ones was performed. The final model performed significantly better than an intercept-only base line model (χ^2(5): 37.227, p = 0) and is a good, optimal fit (C: 0.8354324, Somers' Dxy: 0.670864). The model's explanatory power is moderate (Nagelkerke's R2 = 0.21). No significant results were found, this is likely due to the overall low number of RFR accents.

[8]The model's intercept, corresponding to constituent = First DP and assertion = ExplicitTrueAlternative, is at -1.68 (95% CI [-2.05, -1.35], p < .001).

[9]The model's intercept, corresponding to constituent = First DP and assertion = ExplicitTrueAlternative, is at -1.25 (95% CI [-1.54, -0.97], p < .001).

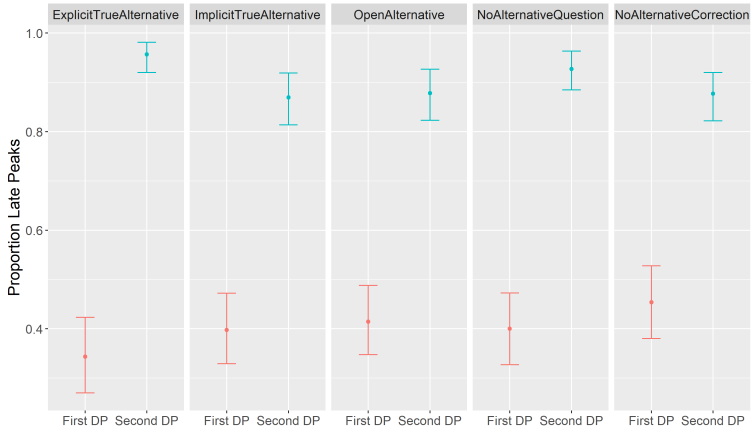


Fig. 6.17: Distribution of late peak accents in German

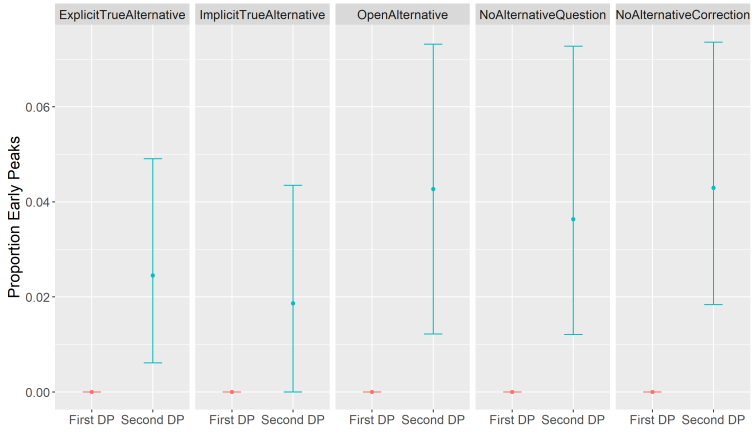


Fig. 6.18: Distribution of early peak accents in German

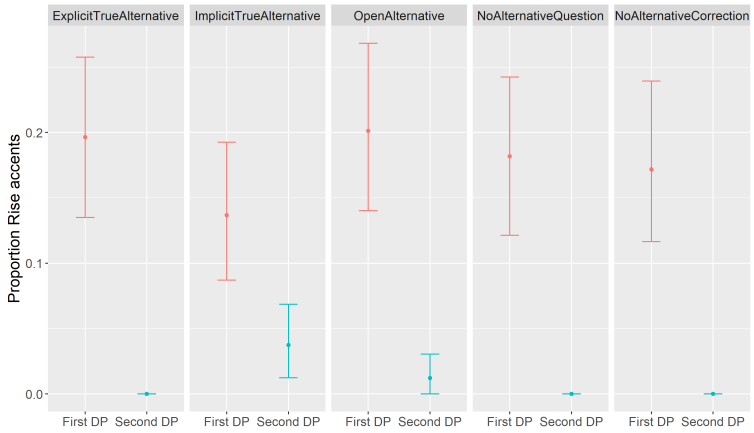


Fig. 6.19: Distribution of Rising accents in German

Fig. 6.20: Distribution of Early Rising accents in German

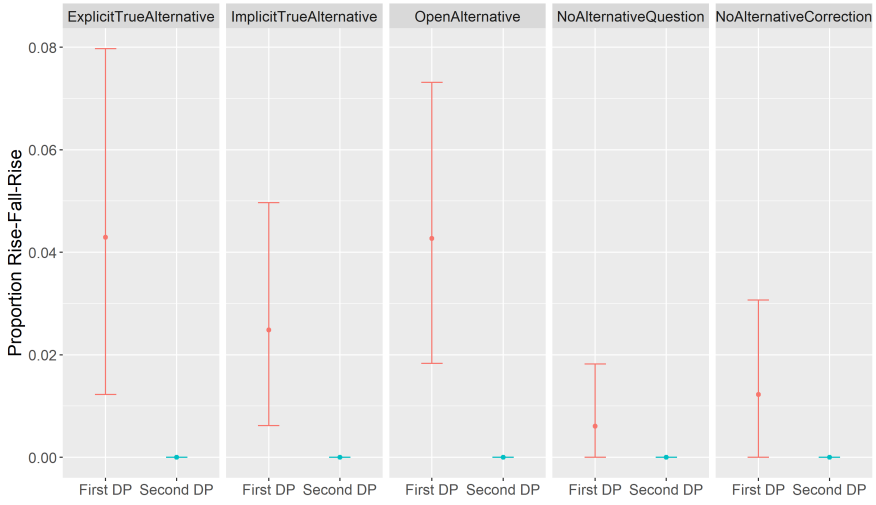


Fig. 6.21: Distribution of Rise-Fall-Rise accents in German

6.6 Conclusion & Discussion

What is immediately evident from the results is that the type of pitch accent is hugely dependent on the linear order in which the pitch accent occurs: Almost only late peaks on the second DP whereas various different pitch accents can occur on the first DP. This pattern is something that has been observed throughout the different languages and experiments, and is most compatible with a Simple-Focus account.

The results do not seem to support any specific account. Wagner (2012) would have expected most hat patterns with an early peak to be found in contexts with true alternatives be it explicit or implicit and Büring (1997) would have expected most hat patterns in contexts with open alternatives. None of these expectations have been found in the data. Interestingly for Dutch the few early peak accents that have been produced seem to be least compatible with the *NoAlternativeCorrection* condition and the same pattern is found for the rising accents. The fact that early peak accents and rising accents display similar patterns was also observed for German in the previous chapter, this again seems to confirm the idea that it is not the individual pitch accents by themselves that map to specific meanings but the tune as a whole. Of course, one should be cautious in drawing any conclusions since there were few data points to analyze and no significant results were reported for the alternative manipulation.

Something else worth observing is the distribution of the RFR accent on the first constituent in Dutch. Although less frequently produced, the RFR accents in Dutch follows more or less the same pattern as in English, i.e. most RFR accents on the first constituent can be found in contexts with explicit true alternatives.

In contrast to Dutch, German does not display any clear pattern in terms of pitch accents other than that the linear order in which the pitch accents occur seem to matter, i.e. almost only exclusively late peaks on the second DP and more options for the first DP. In addition, German is different to Dutch in that there are much fewer RFR accents produced. Interestingly, when comparing these three West-Germanic languages it almost seems as if

Dutch is in between German and English, that is in terms of RFR accents Dutch is more similar to English than German.

In conclusion, Wagner (2012)'s claim that the early peak hat pattern would be most compatible in contexts with true alternative propositions is not supported by the data in this chapter, the hat pattern is used even in contexts in which no alternative can be presupposed to be true. Moreover, Büring (1997)'s claim that the hat pattern comes with an disputability implicature and would therefore be most compatible in contexts with open alternatives was also not supported by the data. All one can say is that in a hat pattern, the rise on the first constituent and the early/late fall on the second simply convey that the sentence has two accented constituents, and this intonation pattern is compatible with double focus contexts. In that sense, the results are similar to BA in English. The results in this chapter were rather weak and only few actual early peak accents were produced. In an attempt to get clearer results, an acceptability experiment was performed in the next chapter specifically testing whether there is a correlation between the different types of hat patterns and the existence of alternative propositions.

Appendix A: Number of Pitch Accents[10]

	Explicit		Implicit		Open	
	First DP	Second DP	First DP	Second DP	First DP	Second DP
Late Peak	55	138	45	133	50	143
Early Peak	0	5	0	12	0	6
Early Rise	52	3	68	1	57	1
Rise	21	2	26	2	32	1
RFR	20	0	9	0	12	0

Table 6.2: Number of pitch accents across different assertion conditions in Dutch

	NoAltQuestion		NoAltCorrection	
	First DP	Second DP	First DP	Second DP
Late Peak	41	125	60	146
Early Peak	0	11	0	1
Early Rise	58	0	70	1
Rise	31	3	15	0
RFR	9	0	3	0

Table 6.3: Number of pitch accents across different assertion conditions in Dutch

	Explicit		Implicit		Open	
	First DP	Second DP	First DP	Second DP	First DP	Second DP
Late Peak	55	136	64	123	66	141
Early Peak	0	4	0	3	0	7
Early Rise	46	0	54	3	46	2
Rise	32	0	20	4	33	2
RFR	7	0	4	0	7	0

Table 6.4: Number of pitch accents across different assertion conditions in German

	NoAltQuestion		NoAltCorrection	
	First DP	Second DP	First DP	Second DP
Late Peak	60	126	63	138
Early Peak	0	6	0	7
Early Rise	41	0	52	0
Rise	30	0	28	0
RFR	1	0	2	0

Table 6.5: Number of pitch accents across different assertion conditions in German

[10] **Go to the OSF-page to access all collected data. Click here.**

Appendix B: Statistical Models

	Late Peak
(Intercept)	$-1.16\ (0.09)^{***}$
constituentSecond DP	$1.09\ (0.07)^{***}$
assertionImplicitTrueAlternative	$-0.15\ (0.10)$
assertionOpenAlternative	$-0.06\ (0.10)$
assertionNoAlternativeQuestion	$-0.12\ (0.10)$
assertionNoAlternativeCorrection	$0.05\ (0.10)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.6: Statistical Models for Dutch: Late Peak

	Early Peak
(Intercept)	$-21.56\ (848.41)$
constituentSecond DP	$18.04\ (848.41)$
assertionImplicitTrueAlternative	$0.85\ (0.53)$
assertionOpenAlternative	$0.15\ (0.61)$
assertionNoAlternativeQuestion	$0.76\ (0.54)$
assertionNoAlternativeCorrection	$-1.64\ (1.10)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.7: Statistical Models for Dutch: Early Peak

	Rise accent
(Intercept)	$-2.05\ (0.21)^{***}$
constituentSecond DP	$-2.78\ (0.36)^{***}$
assertionImplicitTrueAlternative	$0.24\ (0.28)$
assertionOpenAlternative	$0.36\ (0.27)$
assertionNoAlternativeQuestion	$0.36\ (0.27)$
assertionNoAlternativeCorrection	$-0.39\ (0.33)$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.8: Statistical Models for Dutch: Rise

	Early Rise accent
(Intercept)	-1.08 $(0.13)^{***}$
constituentSecond DP	-4.05 $(0.41)^{***}$
assertionImplicitTrueAlternative	0.34 $(0.17)^{*}$
assertionOpenAlternative	0.14 (0.18)
assertionNoAlternativeQuestion	0.11 (0.18)
assertionNoAlternativeCorrection	0.21 (0.18)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.9: Statistical Models for Dutch: Early Rise

	Rise-Fall-Rise accent
(Intercept)	-2.13 $(0.22)^{***}$
constituentSecond DP	-18.46 (851.36)
assertionImplicitTrueAlternative	-0.83 $(0.40)^{*}$
assertionOpenAlternative	-0.54 (0.37)
assertionNoAlternativeQuestion	-0.83 $(0.40)^{*}$
assertionNoAlternativeCorrection	-1.93 $(0.62)^{**}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.10: Statistical Models for Dutch: RFR accent

	Late Peak
(Intercept)	-0.91 $(0.08)^{***}$
constituentSecond DP	0.81 $(0.07)^{***}$
assertionImplicitTrueAlternative	-0.03 (0.10)
assertionOpenAlternative	-0.01 (0.10)
assertionNoAlternativeQuestion	0.02 (0.10)
assertionNoAlternativeCorrection	0.02 (0.10)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.11: Statistical Models for German: Late Peak

	Early Peak
(Intercept)	-21.59 (891.42)
constituentSecond DP	17.88 (891.42)
assertionImplicitTrueAlternative	-0.28 (0.76)
assertionOpenAlternative	0.55 (0.63)
assertionNoAlternativeQuestion	0.39 (0.65)
assertionNoAlternativeCorrection	0.56 (0.63)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.12: Statistical Models for German: Early Peak

	Rise accent
(Intercept)	-1.68 $(0.18)^{***}$
constituentSecond DP	-2.90 $(0.36)^{***}$
assertionImplicitTrueAlternative	-0.12 (0.26)
assertionOpenAlternative	0.08 (0.24)
assertionNoAlternativeQuestion	-0.08 (0.25)
assertionNoAlternativeCorrection	-0.13 (0.26)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.13: Statistical Models for German: Rise

	Early Rise accent
(Intercept)	-1.25 $(0.14)^{***}$
constituentSecond DP	-3.75 $(0.41)^{***}$
assertionImplicitTrueAlternative	0.30 (0.19)
assertionOpenAlternative	0.13 (0.20)
assertionNoAlternativeQuestion	-0.17 (0.21)
assertionNoAlternativeCorrection	0.12 (0.20)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.14: Statistical Models for German: Early Rise

	Rise-Fall-Rise accent
(Intercept)	-3.15 $(0.38)^{***}$
constituentSecond DP	-18.59 (1443.62)
assertionImplicitTrueAlternative	-0.55 (0.63)
assertionOpenAlternative	-0.01 (0.53)
assertionNoAlternativeQuestion	-1.96 (1.07)
assertionNoAlternativeCorrection	-1.25 (0.80)

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6.15: Statistical Models for German: RFR accent

Chapter 7

Hat Contour: Perception Experiment

In the previous chapter it was observed that at least for Dutch both the rising accent and early peak accent displayed similar patterns when presented with contexts with different alternative proposition structures: Both pitch accents were found to be less compatible in contexts without any alternative propositions. The results however, were rather weak and only few actual early peak accents were produced. In an attempt to get clearer results, an acceptability experiment was performed specifically testing whether there is a correlation between the different types of hat patterns, i.e. differing in onset and offset, and the existence of alternative propositions. Based on Wagner (2012), in this chapter it will be argued that only hat patterns with an early peak are less compatible with contexts in which no alternative propositions are present.

7.1 Initial rise and topicality

As discussed in previous chapters, Féry (1989) recognizes two different hat patterns in German: one which consists of an initial early rise (H*) followed by a fall (H*L) (refered to as hat pattern 1) and one which starts of with a rise (L*H) followed by a fall (H*L) (referred to as hat pattern 2). Besides a difference in their phonological form, the two hat patterns are said to differ in meaning as well. Whereas the rise in hat pattern 2 is associated with

topicality, the high plateau of hat pattern 1 is not associated with any special meaning and can be used when no emphasis is needed. As we have seen in previous chapters there is some evidence linking the hat pattern with topicality, although there was no direct evidence that it is the rising accent specifically that is associated with topics.

7.2 Contrastive Topics

Büring (1999; 1997) takes the hat pattern to consist of two phonological units: the initial rise (R) and the final fall (F). For Büring (1997) each phonological element has its own pragmatic/semantic function, the rise constitutes the sentence internal topic (S-topic) and the fall constitutes the focus of the sentence. Whichever element in the sentence receives focus will be part of the set of propositions that are considered well-formed alternatives, this Büring calls the *Focus value*.

S-topics are similar to focus in the sense that they both induce alternatives. These alternatives however, do not have any impact on the focus value, instead, it is a set containing different focus values called the topic value. According to Büring (1997), the topic is indicated by the initial rise in a hat pattern. S-topics come with a certain implicature, namely a disputability implicature which states that there must be at least one element (a set of propositions, i.e., a question) that is still disputable in the context.

7.3 True Alternative Propositions

For Ludwig (2006), the German hat pattern can be interpreted as an operator which takes two foci as its argument and means something like "*there is at least one true proposition (sentence) that is the result of replacing both foci with respective alternatives*". Elaborating on Ludwig (2006), Wagner (2008; 2012) argues that the hat pattern presupposes that an alternative proposition is true. More specifically, it is only the hat pattern with an early peak

(called early peak here) that conveys that a contextually alternative proposition is true besides the proposition expressed bearing the hat pattern.

It follows from Wagner's and Ludwig's definition of the hat pattern that it is less acceptable when there are no alternative propositions. Example (67) illustrates their point: In the example below, the response excludes any alternatives and is therefore rendered less acceptable with a hat pattern. A negative answer conveys that in case of a question with only two options, one option has been excluded and therefore there are no alternative propositions true.

(66) Either Hans insulted Pia or Pia insulted Hans. Did Hans insult Pia?

(67) #Nein. PIA hat HANS beleidigt.
No. Pia has Hans insulted.
'No, Pia insulted Hans.'

The stimuli in the current research are based on this premise, namely that in certain contexts alternative propositions are excluded (e.g. Example 67), but in other contexts such alternative propositions are present. An example of such a context in which an alternative proposition is present is given below (69):

(68) Who insulted who? Did Hans insult Pia?

(69) Ja, und PIA hat HANS beleidigt.
Yes, and Pia has Hans insulted.
'Yes, and PIA insulted HANS.

The hypothesis that will be tested in this chapter is Wagner (2012)'s claim that the early peak hat pattern conveys that a contextually alternative proposition is true besides the proposition expressed bearing the hat pattern.

Example (69) poses a problem for Büring (1997)'s disputability implicature. In contrast to what would have been predicted by this implicature, Example (69) should still contain an unresolved question, however intuitively no such disputability is present in Example (69). According to the disputability implicature Example (69) should imply that besides

Pia insulting Hans, there is still another *unresolved* but related event, which is not the case since everything is resolved. It is true though that a hat pattern in Example (69) does imply that another related event is true, namely *Hans insulted Pia*, but this event is resolved by the affirmative *Ja*. Different from Büring (1997), Wagner (2012) and Krifka (1998) would argue that Example (69) is acceptable since the hat pattern can be used as the last answer to a pair-list question. This follows from Wagner's analysis according to which the hat pattern conveys that there is an alternative proposition which is true. Crucially, Wagner's account differs from Büring's account in that there is no mentioning of any disputability associated with the alternative proposition. The current research will test whether a hat pattern is acceptable when it is the last answer of a pair-list question as in Example (69) and whether the disputability implicature does not hold up in this case. According to Krifka (1998), the first element in a pair-list answer is a contrastive topic and evokes a disputability as argued by Büring (1997), the second element though no longer contains a contrastive topic as it is an answer updated with the first element of the pair-list: the two elements of the pair-list answer together form a complete answer to the question and no disputability is left. Therefore the hat pattern can be used even in cases where there are no contrastive topics.

7.3.1 Expectations

Based on Ludwig (2006) and Wagner (2012), the following hypothesis was formulated: The hat pattern with an early peak conveys that a contextually alternative proposition is true besides the proposition expressed bearing the hat pattern.

It follows from this hypothesis that a hat pattern with a late peak does not specifically express the existence of an alternative proposition. The current study will test this idea by comparing a late peak versus an early peak contour. For the initial part of the contour, two different contours will be examined, one which starts with a rise and one which starts with an early rise. These contours are said to be contrastive with regard to topicality as

discussed earlier by Caspers (1999) and Féry (1989), but they should not be influenced by the existence of alternative propositions. Given this claim it is expected that no difference should be observed in the initial part of the hat pattern in terms the availability of alternative propositions. In order to test for the existence of an alternative proposition, each contour will be embedded in an affirmative or negative answer. The main idea is that an affirmative answer is compatible with an alternative proposition, whereas a negative answer excludes any alternative propositions. A more detailed explanation will be given in the methodology section.

Having presented and explained the hypothesis, one can now make some predictions for the results that follow from these hypotheses. First of all, it is expected that an early peak is less acceptable when embedded in a negative answer which excludes any alternative propositions, than when embedded in an affirmative answer. On the other hand, the late peak should be equally acceptable in both contexts. For the two contour beginnings, the two contours are expected to be equally acceptable in both contexts since they are not influenced by the existence of alternative propositions.

7.4 Methodology

In order to test the hypothesis, an online experiment was set up in the form of an acceptability judgment task using jsPsych (de Leeuw 2015). The participants were presented with a number of carefully manipulated auditory stimuli, after which they were asked to score the stimuli on a scale from 1 to 8 (in which 1 meant completely unnatural and 8 meant completely natural).

7.4.1 Participants

A total of 60 participants were recruited, 30 of them were native speakers of Standard Dutch (as spoken in the Netherlands) and the other 30 were native speakers of Standard

German (as spoken in Germany). A compensation of 2.33 C$ was given as compensation which boils down to an hourly rate of 13.98 C$ since the experiment itself took about 10 minutes.

7.4.2 Stimuli

Each stimuli set consists of three consecutive elements: (1) A context, which introduces the conversation topic, (2) a question related to the context, and (3) an answer to the question (see 70). The context is presented to the participant in written form whereas both the question and answer are presented in written and auditory form. A more detailed description of how these elements are presented to the participant is given in the *procedure* section. An example of such a stimuli set is provided below:

(70) **Context:***Willem-Jan was walking home when coincidentally Isabel was also walking home on the other side of the street. Because they were both in a hurry, they almost walked passed each other without noticing one another.*

Question: Who had seen who? Had Isabel seen Willem-Jan?

Ja, en WILLEM-JAN had ISABEL gezien.
Yes, and Willem-Jan had Isabel seen.

Answer: 'Yes and Wilem-Jan had seen Isabel.'

Each question in the experiment is set up in such a way that the subject and object in the question (in this case *Isabel* and *Willem-Jan* respectively) contrast with the subject and object in the answer (in this case *Willem-Jan* and *Isabel* respectively). Thus creating a double contrast which accents the same two elements in the sentence across different stimuli. In addition, no new persons are introduced in the answer, in other words all persons in the target stimuli are given rather than completely new. This is to accommodate the incompatibility of the (early peak) hat pattern with new information as we have seen from the previous literature in the previous chapter. The target stimuli have been manipulated along the following parameters:

Context: embedded within either a Yes or No answer.

Beginning: A rise (L*H) or early rise (H*) initial part of the contour.

Ending: A late peak (H* L%) or early peak (H!H* L%) ending of the contour.

Recall from the introduction that the existence or absence of an alternative proposition is essential to our definition of the hat pattern. Crucially, the idea here is that in case of an affirmative answer, there exists an alternative proposition, whereas in case of a negative answer this alternative proposition is absent. For instance, the affirmative answer in (70) conveys that besides the event *"Willem-Jan had seen Isabel"* another alternative event is true, namely *"Isabel had seen Willem-Jan"*. The negative answer on the other hand (see 71), conveys that only the situation in which *"Willem-Jan had seen Isabel"* is true and that the alternative event *"Isabel had seen Willem-Jan"* is not true. The context for the negative answers differs from the ones for the affirmative answers in the sense that all alternative propositions are excluded. Once again the question is asked *"Who had seen who, Isabel Willem-Jan?"*, but this time the answer is negative. The last sentence of the negative context is crucial, as it indicates that only one of them had seen the other person therefore no other alternative propositions exist (see 71).

(71) **Context:***Willem-Jan was walking back home when coincidentally Isabel was walking on the other side in opposite direction. Because they walked passed each other so quickly they almost didn't notice the other person. Luckily one of them had spotted the other person.*

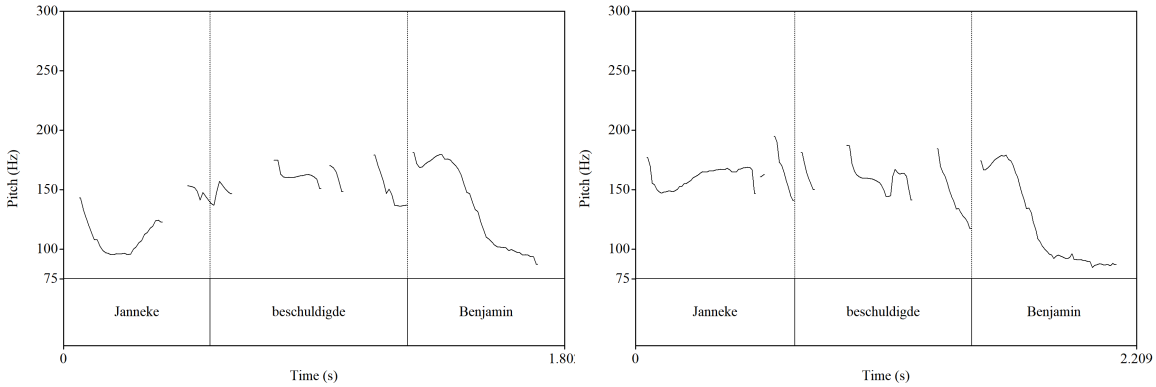
Question: Who had seen who? Had Isabel seen Willem-Jan?

Nee, WILLEM-JAN had ISABEL gezien.
No, Willem-Jan had Isabel seen.

Answer: 'No, Wilem-Jan had seen Isabel.'

In terms of the beginning of the contour, it either starts with an initial rise, as illustrated in Figure 7.1a, or with an initial early rise (see 7.1b).

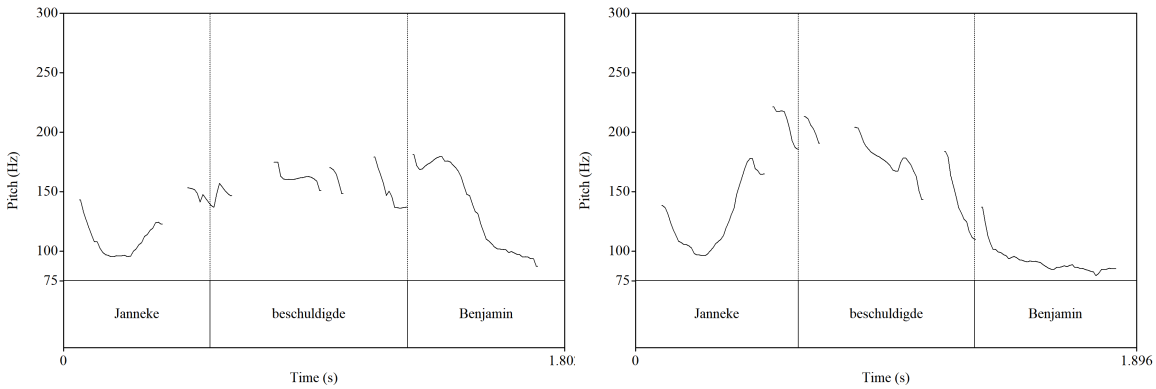
The ending is expected to be heavily influenced by the availability of an alternative proposition or not: having a hat pattern with an early peak is expected to be significantly



(a) Rising accent (b) Early Rise

Fig. 7.1: Example of contour onsets used as auditory stimulus in the experiment

more acceptable when there are alternative propositions than when there are not. Within the late peak contours, no difference in context is expected. Phonetically speaking the late peak occurs within the accented syllable, which is *Benjamin* in Figure (7.2a), whereas the early peak occurs somewhat before the accented syllable, i.e. before *Benjamin* (see 7.2b).



(a) Late peak (b) Larly peak

Fig. 7.2: Example of contour endings used as auditory stimulus in the experiment

The stimuli were recorded in a sound-attenuated booth (16 bit, 44100 Hz) by native speakers of German and Dutch. 16 stimuli sets have been created, there are eight different conditions [1] meaning that a total of 128 stimuli have been recorded (16 sets * 8 conditions = 128). The stimuli are presented to the participants in a latin-square design.

[1] 2 (alternatives/no alternatives) * 2 (L*H or H* beginning) * 2 (H*L or H!H* ending) = 8 conditions

7.4.3 Procedure

Participants are directed to the experiment via an online link. The main experiment starts off with a fixation point, each trial starts off with a fixation point, this is to ensure the participant knows exactly when one trial has ended and another one starts. Next, the context is given in written form: the context ensures the phrases sound more natural rather than out of the blue (see the stimuli section for an actual example).

After that, a screen appears indicating that a question is coming up (indicated in red) together with a image showing what has happened (7.3). At the beginning of the experiment it is explained that the image indicates the actual situation in the real world. This image, depicting the current situation, remains on screen throughout the rest of the trial. The arrows indicate who did what to whom, in this case one arrow goes from Willem-Jan to Isabel thus indicating that Willem-Jan inflicts an action on Isabel and another arrow the other direction thus indicating that the action is reciprocal. Above the arrow, the specific action is written which in this case is *gezien* "seen". So the image indicates that two events have happened in the actual world: (1) Isabel had seen Willem-Jan, and (2) Willem-Jan had seen Isabel.

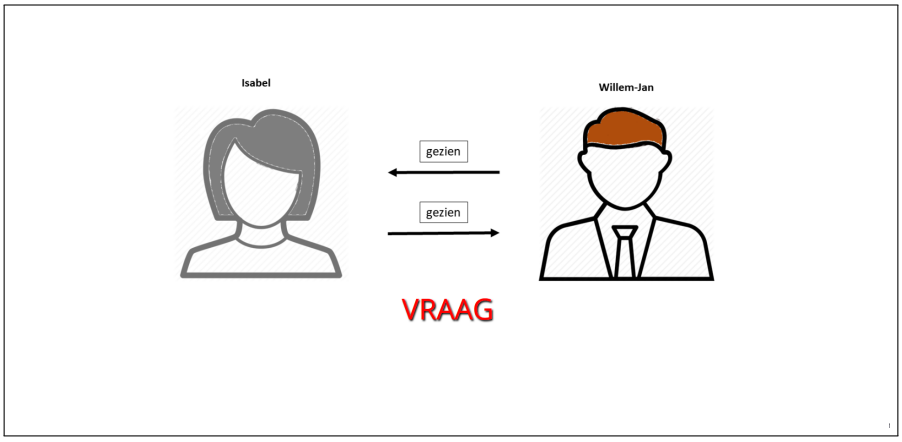


Fig. 7.3: Question

After 1500 milliseconds the participant will hear the question, which could be for example: *"Who had seen who, Isabel Willem-Jan?"*. All questions in this experiment contain

multiple wh-words, this enforces the use of two foci in the answer. By providing the subject and object in the question we have introduced these persons and from now on they will be part of the given information structure.

The question trial disappears and the answer trial appears which has a similar layout but now the word 'answer' is written in red. All trials with alternative propositions contain images like the one in Figure (7.3), i.e. in which there is more than one event. In the *no alternatives* condition, on the other hand, there is only one arrow, i.e. only one event (see 7.4). The image in Figure (7.4) enforces the idea of no alternative propositions by showing that there was only one seeing event, namely one in which Willem-Jan had seen Isabel.

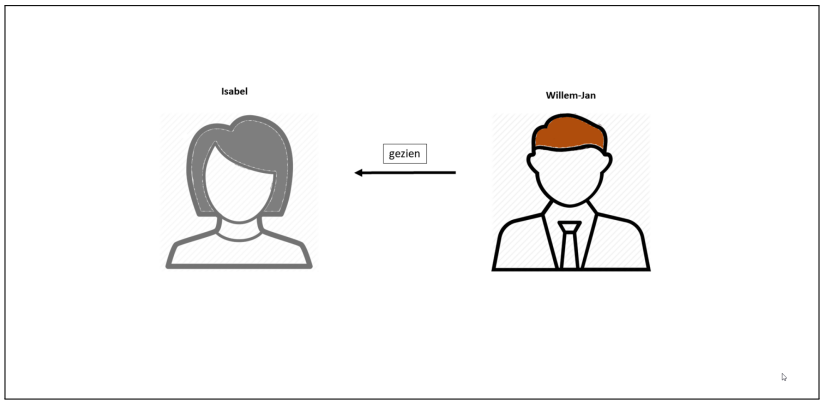


Fig. 7.4: Negative question

Once again a sound is played with the target stimuli, which in case of an affirmative answer could be: *"Yes and Willem-Jan had seen Isabel."*. The affirmative answer implies the existence of alternative propositions: Besides the event that Isabel had seen Willem-Jan, there was another event in which Willem-Jan had seen Isabel. Remember that we created 4 different contours, since in this case the situation does not exclude any alternative proposition, we expect the early peak contour to be assessed relatively acceptable in this context. Also, keep in mind that there is a double contrast between the question and answer pair: Whereas in the question *Isabel* is the subject and *Willem-Jan* the object of the sentence, the reversed is true for the answer thus creating a double contrast. In addition,

no new persons are introduced between the question and answer, therefore all persons in the answer are given.

Finally, the trial finishes off with an acceptability input screen in which the participant has to indicate how natural the intonation pattern sounds given the answer, based on a scale from 1 to 8 where 1 means completely unnatural and 8 means completely natural.

7.5 Results

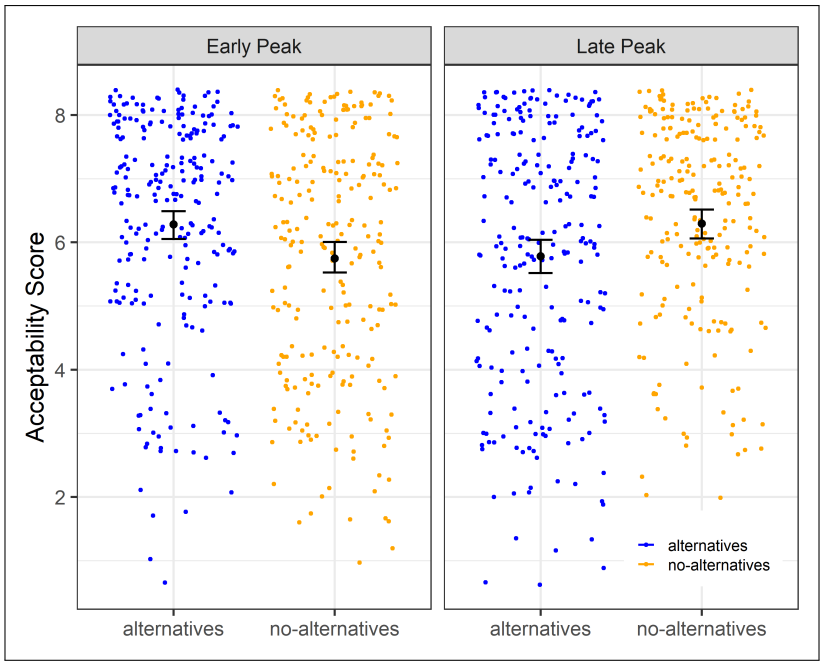


Fig. 7.5: Acceptability score per context, faceted by ending.

As predicted by our hypothesis there seems to be a clear contrast between having alternative propositions or not with regard to the early peak: The existence of alternative propositions is more acceptable than the absence of it. Interestingly, the opposite pattern is observed for the late peak, i.e. the *alternatives* condition is less acceptable than the *no alternatives* condition. This is not per se what was expected, but it also does not contradict the hypothesis.

To see whether the observations from the figure are significant cummulative link mixed models analysis (CLMM) was run using R (Team 2008) and the ordinal Rpackage (Christensen 2019). The outcome variable was the acceptability score on a likert scale from 1 to 8. As fixed effects *context*, *contour ending* and *contour beginning* as well as their interactions were added to the model. As random effects, by-subject random intercepts were included in the model.[2] From the model comparisons it became clear that only the interaction between context and contour ending was significant.

A model with only *contour ending* (AIC 3421.4) did not significantly improve over an intercept-only null model (AIC 3419.91) (χ^2 (1, N = 960) = 0.472, p = 0.492). Adding *context* as a fixed effect also did not improve the model (AIC 3423.4) (χ^2 (1, N = 960) = 3e-04, p = 0.986). However adding an interaction effect between *context* and *contour ending* did significantly improve the model (AIC 3404.0) (χ^2 (1, N = 960) = 21.431, p = 3.668e-06).

Adding language to the model (AIC 3405.2) did not significantly improve the model (χ^2 (1, N = 960) = 0.763, p = 0.382) nor did a three-way interaction between language, context and contour ending (AIC 3404.8) improve the model (χ^2 (3, N = 960) = 6.481, p = 0.0904).

As for the contour beginnings, the expectation was that there would be no difference across the different contexts. Contrary to our expectations though, there does seem to be a difference according to which rising accents are more acceptable with alternative context than high accents, and vice versa (see 7.6. From the model comparison it becomes clear that adding an interaction effect between *contour beginning* and *context* (AIC 3382.2) did significantly improve the model compared to a model without such interaction effect (AIC 3398.2) (χ^2 (1, N = 960) = 18, p = 2.194e-05).

Similarly to the contour endings, language did not have a significant effect: Adding a three-way interaction between language, context and beginning to the model (AIC 3383.9) did not significantly improve the model (χ^2 (1, N = 960) = 4.2372, p = 0.237).

[2]I will refer to this model as Model 1 henceforth, see Appendix A for all statistical models.

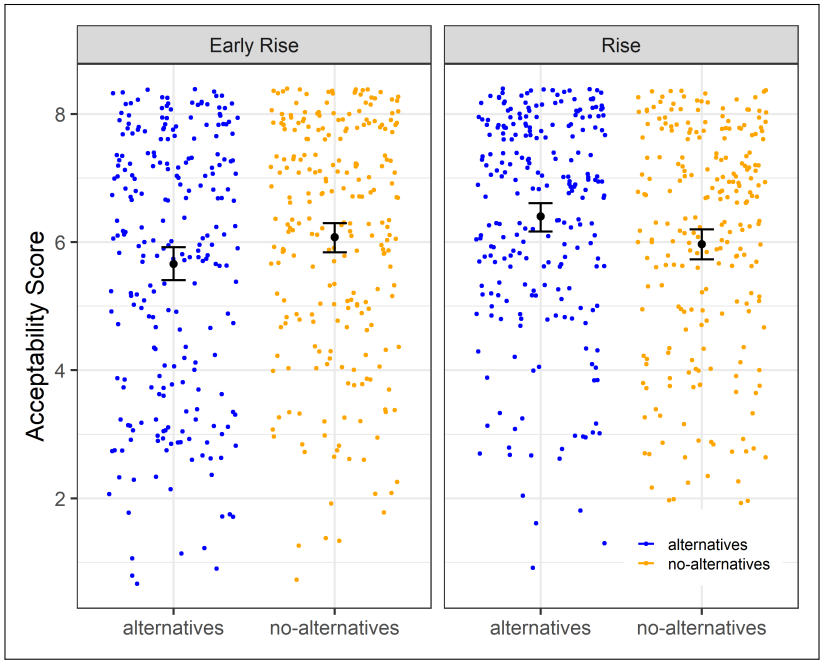


Fig. 7.6: Acceptability score per context, faceted by beginning.

7.6 Conclusion

From the results it can be concluded, that the contour endings are significantly different across the different contexts. This supports our hypothesis which claims that there would be a significant difference between having alternative propositions or not with regard to the ending of the hat pattern. Moreover, the hypothesis claims that an early peak should be significantly less acceptable when no alternative propositions are true than when there are such alternatives. As can be seen from Figure 7.5, within the early peak the *alternatives* condition is more acceptable than the *no alternatives*. For the late peak contours the opposite pattern is observed.

So far the results for the endings of the hat pattern seem to support the hypothesis. However, it is important to note that the effects are rather small (which is not uncommon for prosodic differences): the difference between the *alternatives* and *no alternatives* conditions seems to be no more than 0.5 on a Likert scale from 1-8. This could either be because the hypothesis describes a trend rather than a categorical effect of the meaning of the contour, or it could be that listeners are just easily willing to disregard prosody.

Notice also that even though it was predicted that the *no alternatives* condition would be significantly worse than the *alternatives* condition for contours with an early peak, the *no alternatives* condition is relatively acceptable. In fact, it is as acceptable as the alternative context with a late peak. Once again, this tells us that one is dealing with a trend rather than a categorical difference.

With respect to the contour beginnings, the prediction was that there would be no difference across the two context conditions. However the results tell us that there is a significant interaction effect between contour beginning and context: The rising accent is preferred over a high accent in contexts with alternatives, and vice versa. In a way, the rising accent behaves similar to the early peak accent.

7.7 Discussion

The current research tested the hypothesis that the early peak in a hat pattern indicates that an alternative proposition must be true besides the proposition expressed bearing the hat pattern. This hypothesis is supported by the results of our online experiment: For the early peak, the *alternatives* condition is significantly more acceptable than the *no-alternatives condition*. I have thus found evidence supporting Ludwig (2006) and Wagner (2012). However, one has to keep in mind that the effects were rather small: The difference between the *alternatives* and *no alternatives* condition for the early peak was around 0.5 points on a 8-point scale. We are thus dealing with a trend rather than a categorical distinction. In addition, the results from this research do not support Büring's disputability implicature: Even though there was arguably no disputability in the *alternatives* condition, the hat pattern was still acceptable (with hat pattern either the early peak or late peak hat pattern is meant in this case).

With regard to the contour onset, our predictions were that there would be no difference for the initial part of the hat pattern in terms of the existence of alternative proposi-

tions. However, there was a significant difference across different context conditions for the hat patterns with an initial rise: The *alternatives* condition was significantly more acceptable than the *no alternatives* condition. It thus seems that both the rise accent and the early peak accent are more acceptable in the same context. This is in line with the results of other experiments in this study where the conclusion was drawn that the intonational tune as a whole conveys a certain meaning rather than the individual pitch accents.

In conclusion, both the early peak and rise accent are less acceptable in contexts without alternative propositions. The fact that similar results were obtained in production experiments as was tested in the previous chapter (albeit non significant), reinforces this idea. Turning our attention to our main research question of how intonational form and meaning are mapped, one can conclude that there is no direct mapping between a pitch accent and a specific meaning as was claimed by Jackendoff (1972) and Büring (1997). Rather it is the intonation pattern as a whole which maps to certain meanings. For the hat pattern specifically the conclusion can be drawn that the hat pattern does not convey which constituent is a contrastive topic or a focus, instead it can be used just convey that two constituents evoke alternatives.

Appendix A: Statistical Models

	Cummulative Link model
ending	$0.08\ (0.12)$
context	$0.01\ (0.12)$
language	$0.26\ (0.30)$
beginning	$0.36\ (0.12)^{**}$
ending:context	$1.10\ (0.24)^{***}$
context:beginning	$-1.00\ (0.24)^{***}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 7.1: Statistical Model

Chapter 8

Conclusion and Discussion

At the very start of this thesis the question was asked whether it is the full intonational tune that conveys a certain meaning (holistic approach) or the decomposed contours that make up an intonational tune (compositional approach). Throughout the various experiments we have seen that, contrary to claims made by several studies (Jackendoff 1972; Büring 1997), there is no evidence supporting the idea that specific pitch accents rather than the overall pattern convey certain meanings (at least for AB-BA constructions in English and the hat pattern in German and Dutch). In this chapter we will readdress the research questions posited in the first chapter and discuss them in more detail.

8.1 Meaningful tunes vs. contours

Let us first delve into the question of whether intonational meaning is best analyzed holistically or compositionally. Throughout this thesis we have seen evidence for a more holistic approach to intonation rather than a compositional approach: There is a holistic tune conveying a meaning rather than separate independent contours each carrying their own semantics. In other words, the holistic tune constitutes an intonational morpheme rather than the decomposed contours that make up an intonational tune. Even though the ABBA patterns and hat patterns seem to be best approached from a holistic point of view, that

is not to say that intonation in general should be analyzed as holistic. It could very well be that other intonation patterns are best analyzed compositionally, all this research has shown is that a compositional approach to the ABBA and hat patterns is not supported.

8.1.1 AB-BA Patterns

This thesis was split up in two parts, namely the AB-BA patterns in English and the hat patterns in German and Dutch. I will first discuss the results from English before turning my attention to German and Dutch. Jackendoff (1972), Büring (1997) and Constant (2014) hypothesized that there is a direct connection between contrastive topics/foci and specific contours[1] in English: B accents mark a contrastive topic and A accents mark a focus. The results from the experiment in chapter 2 showed that first of all Jackendoff (1972)'s B accent (i.e. Rise-Fall-Rise accent) was rarely produced. Second and most importantly, no consistent correspondence was found between an accent and the information status of a constituent (i.e. contrastive topic or focus): This became clear both from the fact that different focus conditions did not display different tunes, nor did syntax have an influence on the tune (i.e. the active and passive sentences were produced with the same intonation pattern even though the position of CT and Focus would have switched around). Instead, the type of contour turned out to be highly dependent on whether the contour comes initially or finally, with a high proportion of B accents on a DP if a second accented DP followed. Therefore, one can conclude that B and A accents do not mark whether something is a (contrastive) topic or focus, but rather, using a B accent encodes that another accented DP follows.

A different approach to the AB-BA patterns was taken by Wagner (2012). According to Wagner (2012) the BA and AB patterns are fundamentally two unrelated intonation patterns each with their own semantics. According to this analysis, the BA pattern is simply the declarative fall, where non-final accentuated constituents can be realized with

[1]Jackendoff (1972) refers to it as "pitch accents" even though from his description of the "pitch accents" (i.e. pitch accent plus boundary tones) it becomes clear the term *contour* is more appropriate.

a rise. This is compatible with the prediction of the experiments in Chapter 2. The AB pattern indicates that an alternative proposition is potentially true. This is in line with the results in Chapter 2 in that the AB contour was almost never used. This is expected given that these dialogues did not suggest that the respondent necessarily wanted to convey that there are still open alternatives.

The experiment in Chapter 3 looked at the predictions of the various theories in a more fine-grained way. The results from Chapter 3 only partially supported Wagner (2012)'s predictions: BA patterns were most frequently produced in contexts with explicit alternative propositions, this is compatible with the idea that BA patterns require a contextually contrastive item. However, it should be noted that this is not what Wagner (2012) would have expected. According to Wagner (2012), the B accent in BA patterns is completely optional. The BA pattern then is simply a falling declarative contour, where non-final accented constituents can be optionally realized with a B accent and final accented constituents with an A accent. In other words, the use of B and A accents is completely determined syntagmatically by the position in an utterance, and context does not have any meaningful influence on the contour. This is indeed what was found in Chapter 2, in Chapter 3 however it becomes clear that BA patterns, as opposed to AA patterns for instance, are specifically more appropriate in contexts with true alternative propositions: Besides an effect of linear order as observed in Chapter 2, the results from Chapter 3 show a clear effect of context as well.

As for the AB patterns, few actual AB patterns were produced in general and therefore no conclusions could be drawn for AB patterns. Despite the fact that the expectations were only partially supported by the results, the results do provide evidence supporting a holistic approach to intonational meaning: BA patterns as a whole are more appropriate in contexts with explicit alternative propositions, there is however, no specific meaning that can be assigned to the B and A accents respectively.

Büring (1997)'s claim that AB-BA patterns come with a disputability implicature were not supported by the results from Chapter 3: Given this claim the expectation would have been that AB-BA patterns are most compatible in contexts with open alternatives, no such result was found. Previous research however, has found the AB pattern to be most appropriate in contexts with incomplete answers which tells us that the AB patterns convey some kind of disputability implicature is not completely wrong, but Büring (1997)'s semantic analysis of this implicature was not supported by the results of this study. The results also do not support Jackendoff (1972)'s analysis who predicted no AA patterns given that there can be only one primary focus in a sentence.

All in all one can conclude that the experiment in Chapter 3 is in line with the results from the experiment in Chapter 2 in that neither of the experiments seem to support the idea that there is an exact correspondence between a contour (i.e. decomposed part of a tune) and a specific meaning associated with that contour. It should be noted however that even though the ABBA patterns seem to be best approached from a holistic point of view, that is not to say that all intonation patterns should be analyzed as purely holistic.

To conclude this section on English, let us summarize where the Jackendoff (1972) paradigm went wrong. First of all, the BA contour does not convey whether the B-accented or the A-accented constituent functions as a contrastive topic in the utterance, it simply conveys that two constituents evoke alternatives, and non-final ones can be realized with a rise: This is exactly what was expected by Wagner (2012). The second basic insight that Jackendoff (1972) was wrong about is the claim that the AB pattern is merely the mirror image of the BA pattern, the results showed us that the AB pattern is rarely used and if used conveys a meaning different from what the BA pattern conveys. Wagner (2012)'s nested focus account also made wrong predictions as we have seen from the data: The word order for sentences with multiple foci that were claimed to be infelicitous (i.e. sentences in which the accented DP contextually serves as a contrastive topic follows the accented NP which contextually serves as the focus) are in fact felicitous, and the AB pat-

tern (or Rise-Fall-Rise pattern) is not likely to be used in contexts in which an alternative is possibly true, in reality it is rarely produced across the board.

In the end, the Simple-Focus account, according to which non-final accents can be optionally produced with a rising accent, seems to best capture the results for English: According to this account the type of accent is mainly determined by the linear order in which an accent occurs. There are however additional factors that must play a role in the phonetic realization of accents, e.g. think of BA patterns being more frequently produced in contexts with explicit true alternatives.

8.1.2 Hat Patterns

Chapter 4 explored different phonological representations for the early and late peak accents which form a crucial part of the hat pattern. Since there is no consensus on the "correct" phonological representations of these two falling accents and especially the early fall accent, several experiments were performed to explore the differences between the two falling accents. From the phonetic results it was concluded that the early peak accent is best represented as "(%H) H+!H* (L%)" and the late peak accent as "H* L%".

Turning our attention to the semantic side of hat patterns, like the English experiment, the results from Chapters 5, 6 and 7 showed that the hat pattern does not convey which constituent is a contrastive topic or focus, instead the hat pattern can be used just to convey that two constituents evoke alternatives.

Much like his compositional approach to English AB-BA patterns, Büring (1997) argues for a similar compositional approach to the hat pattern in German: Rising accents mark a contrastive topic and falling accents mark a focus. The results in Chapter 5 were in some ways similar to the analogous experiment run in Chapter 2. The difference however being that some pitch accents seemed to indeed be more frequently produced on contrastive topics than on foci. There was a tendency of more rising accents on contrastive topics and early peak accents almost exclusively occurred on foci. However, although

rising accents were indeed more frequently produced on contrastive topics, a substantial proportion of rising accents was also produced on constituents that would be considered the focus of the sentence. In addition, the early peak accent also displayed the expected pattern, but very few actual early peak accents were produced. Moreover, a large proportion of contrastive topics were produced with a falling accent, thus providing even more evidence against Büring (1997)'s claim that each pitch accent is assigned a specific meaning.

For Wagner (2012) the hat pattern with an early peak accent as a whole indicates that there is a true alternative proposition. In contrast to Büring (1997), Wagner (2012) does not assign a specific meaning to each individual pitch accent but instead argues that the whole intonation pattern itself conveys the meaning. The claim made by Wagner (2012) is that the early peak hat contour should be most felicitous in contexts with true alternatives. This claim was put to the test in Chapter 6. Contrary to expectations, the hat pattern was not more appropriate in contexts with true alternatives and hat patterns were even produced in contexts without alternative propositions. Instead it was observed that both the rising accent and early peak accents tended to be produced less in contexts with no alternatives. This observation was corroborated by the results from acceptability judgment experiment in Chapter 7: The rising accent as well as the early peak accents were less acceptable in contexts with no alternative propositions compared to contexts with such alternatives.

The results from both Chapter 6 and 7 tell us that rather than pitch accents being independent and having their own semantics, it seems that multiple pitch accents together convey the same semantics and thus there is evidence supporting a more holistic view towards intonational meaning at least for the hat pattern.

Another piece of evidence against a compositional view of the hat pattern comes from the phonological production experiment in Chapter 4. Although not explicitly tested in Chapter 4, some of the intonations produced by participants provide evidence of a more holistic approach to intonation in which a combination of pitch accents give rise to a cer-

tain semantics rather than each pitch accent by itself having its own semantics. If there is truly a direct correspondence between meaning and pitch accent then one would not expect to find two pitch accents within the same word as is the case in some of the productions made by participants.

All in all one can conclude that there is no evidence supporting the claim that intonational meaning is compositional. Although there is some evidence that holistic tunes can be associated to certain meanings and the data seem to be most compatible with the idea that tunes have a constant meaning, this study has not been able to identify an exact meaning yet as the contexts that were tested were compatible with multiple contours. In the following sections we will turn to the second research question, namely how tunes and the metrical structure are build up.

8.2 Metrical Structure and Focus

In this section we address the question of how the metrical structure is realized in terms of accentuation. In the very first chapter it was mentioned that the metrical structure of a sentence is determined by various factors among which focus is one of them and the one that we are most interested in for the purpose of this study. The assumption is that a wh-question determines the focus in a sentence, the focused constituent is then considered metrically prominent and realized with an accent. In Example (72a) for instance the first DP is focused and therefore the expectation would be that this focused DP is accented and is realized with a pitch accent. The non-focused DP on the other hand would not be accented and therefore not receive a pitch accent, this is exactly what was found in the results in this study. In a similar fashion, when both DPs are focused the results show that both DPs are accented and each receive a pitch accent (see 72c). However, when the second DP is focused one does not observe the expected pattern: Even though the metrical structure tells us that there is only one (focus-induced) accented constituent, in

actual productions two constituents are accented and receive a pitch accent (see 72b). The condition with focus on the second DP then seems to behave similar to the double focus condition.

(72) a. *What about the beans? Who ate those?*

[Fred	ate the beans]	*First DP focus*
x		(focus)
accented	deaccented	(accentuation)

b. *What about Fred? What did he eat?*

[Fred	ate the beans]	*Second DP focus*
	x	(focus)
accented	accented	(accentuation)

c. *Who ate what?*

[Fred	ate the beans]	*Double focus*
x	x	(focus)
accented	accented	(accentuation)

In this study we have seen cases of double accentuation in sentences like (72b) as well as (72a) and the idea that such sentences can be optionally realized with both DPs accented when interpreted as a contrastive topic construction is the very premise on which Jackendoff (1972)'s analysis of AB-BA patterns is built. As discussed in chapter 2, to account for such double-accented realizations one could consider contrastive topic constructions to contain an implicit multiple-wh question (see 73), with this implicit question determining the metrical structure. This could explain why double-accented realizations are observed in all focus conditions. However, it cannot explain why there is an asymmetry between the first DP focus and second DP focus conditions in terms of deaccentuations.

(73) *(Who ate what?) What did Fred eat?*

[Fred	ate the beans]	*Second DP focus*
x	x	(focus)
accented	accented	(accentuation)

To explain this asymmetry I will assume that in the second DP focus condition there is indeed only one focus-induced accent in the second DP focus condition (just like in

the first DP focus condition), but the non-focused constituent would then be a case of "accent without focus".[2] In the next section I will argue that this asymmetry arises due to phonological factors.

Interestingly this asymmetry is observed throughout the different experiments and across the three languages. Based on this pervasive result we can conclude that there is an asymmetry in how the metrical structure is realized by pitch accents. In a way this is analogous to something we have seen at the end of Chapter 4 where only one-word answers were recorded: For words with non-initial lexical stress (i.e. words with either penultimate or final stress) some words were produced with two pitch accents even though metrically speaking there was only one stressed syllable. Similar examples of "additional" accents on words consisting of multiple syllables have been noticed in the literature but remain problematic in many intonational theories (see Selkirk (1986) and Gussenhoven (1983)). A possible solution would be to make a distinction between primary (or nuclear) and secondary accents and only consider primary accents to be relevant to signalling focus (Ladd 2008). The idea that the last accent in a certain domain has a special status because it is perceived as the most prominent in a series of equal accents, goes back to Newman (1946).[3]. Secondary accents (or non-nuclear accents) on the other hand are distributed according to other criteria (Ladd 2008), these so-called "other" factors could be phonological factors as we will see in the next section.

8.2.1 Asymmetry between Pre- and Post-nuclear Position

I argue that the asymmetry discussed in the previous section is best captured by the idea that constituents in pre-nuclear position can be optionally accented, while in post-nuclear position such constituents tend to be deaccented: The nuclear accent in this case is determined by the focused constituent as we have seen in the previous section.

[2]see Ladd (2008) Chapter 7 for a discussion on "accent without focus" in broad focus cases.
[3]For a more detailed discussion on nuclear accents and focus see Truckenbrodt (1995) and Wagner (2005)

The double focus construction is the least controversial condition (see 74): Both *Fred* and *beans* are F(ocus)-marked and therefore receive a pitch accent with the last accent being considered the nuclear accent. Unlike the double focus conditions, for the other focus conditions the speaker can optionally choose a double focus interpretation or a single focus interpretation depending on whether the speaker has an implicit multiple wh-question in mind or the immediate wh-question. When opting for a double focus interpretation, both DPs would be F-marked and realized with an accent, thus resulting in a metrical structure (as well as phonetic implementation) indistinguishable from the double focus condition (see Examples 75 and 76). This may lead one to falsely believe that a focused constituent is always realized with a pitch accent and vice versa.

(74) *Who ate what?* (Double Focus)

[Fred]$_F$	ate the	[beans]$_F$
PreNuc	Nuc	(order)
Accent	Accent	(accentuation)

(75) *What about Fred? What did he eat?* (Second DP focus)

[Fred]$_F$	ate the	[beans]$_F$
preNuc	Nuc	(order)
Accent	Accent	(accentuation)

(76) *What about the beans? Who ate those?* (First DP focus)

[Fred]$_F$	ate the	[beans]$_F$
preNuc	Nuc	(order)
Accent	Accent	(accentuation)

In cases with single focus however, the accent realizations do not always reflect the focus structure of a sentence. Indeed for the *First DP focus* condition only F-marked constituents are prominent and receive a pitch accent (see 77) with the non-focused constituent (in this case *beans*) being deaccented. These relate to the jackendoff paradigm as well, in that the Foc-CT order is preferentially realized with deaccentuation, showing that marking both DPs as focused might actually be the exception rather than the rule in this paradigm. On the other hand when the context forces the focus to be on the second DP,

it was observed in this study that non-focused material can still receive a pitch accent (see 78). The explanation I provide for this observation is that unlike pre-nuclear accents, post-nuclear unfocused constituents are realized without accent. This exact asymmetry has been observed by several researchers (Büring 2016; Ladd 2008; Wagner 2005).

(77) *What about the beans? Who ate those?* (First DP focus)

[Fred]$_F$ ate the beans		
Nuc	postNuc	(order)
Accent	Deaccented	(accentuation)

(78) *What about Fred? What did he eat?* (Second DP focus)

Fred	ate the [beans]$_F$	
preNuc	Nuc	(order)
Accent	Accent	(accentuation)

To summarize, *focus* determines the metrical structure of a sentence by assigning prominence to the focused constituent, the prominent constituent is then accented and receives a pitch accent. The last accent in an utterance is considered the nuclear accent. Non-focused constituents can receive a “secondary" accent in pre-nuclear position, but tend to be deaccented in post-nuclear position. Féry and Kügler (2008) studied the influence of information structure on tonal scaling in German and found similar results: Givenness lowers the peaks of rising accents in pre-nuclear position whereas in post-nuclear position constituents are completely deaccented. I will come back to Féry and Kügler (2008) in a later section. A different approach to explain this asymmetry comes from Calhoun (2007; 2012); Bishop (2013). As we have seen in the discussion at the end of Chapter 2, Calhoun (2007; 2012) claims that a theme-rheme distinction in English is cued by pitch range: Themes are lower in pitch range than rhemes. However, Calhoun’s notion of ’theme’ arguably conflates the case of deaccented non-foci and accented non-foci and therefore the results from her study do not convincingly provide evidence in favor of this hypothesis. Nonetheless, even though Calhoun (2012)’s implementation of the theme-rheme distinction is potentially confounded, the idea that pitch scaling rather than pitch

accent plays a role in encoding topics and foci has been argued for by other researchers like Bishop (2013).

A prosodic subordination approach can indeed also account for the results in Chapter 2. However according to my interpretation, pitch scaling is used to encode prominence shifts to mark focus, it does not encode topics and foci specifically as proposed by Calhoun (2012) and Bishop (2013). Especially the *double focus* contexts turn out to be problematic within a prosodic subordination account in which a lower pitch range encodes for focus and a higher pitch range encodes for topics: The double focus contexts do not explicitly give rise to contrastive topics, but the results are very similar to the *second DP focus* contexts (i.e. contexts in which there is arguably a contrastive topic). It is therefore unlikely that pitch range is a salient cue to encoding topic-focus configurations. Instead prosodic subordination is a way to encode what is non-focused in the scope of the squiggle operator.

Regardless of whether or not there is a prosodic subordination that encodes the topic-focus configurations, the main point made in Chapter 2 still stands: B accent and A accents do not mark for topics and foci respectively, thus invalidating Jackendoff's and Büring's claims about contrastive topics. The notion of B accent as being deaccented or lower in pitch range than A accents is not what Jackendoff or Büring had in mind when discussing the B and A accents. In the next section I will turn my attention to how tunes align with the metrical structure.

8.2.2 Metrical Structure and Pitch Accents

How does metrical structure lead to the particular distribution of pitch accents we observe? Based on the results from this study one can draw some concrete conclusions on how the two layers interact. Basically the type of pitch accent depends strongly on how many accented syllables there are in a sentence and whether the pitch accent is final or not. For non-final accents a whole range of different pitch accents can be produced while

the final accented constituent has a much more restricted set of pitch accents that can be produced: For all three languages tested in this study the final accent is mainly a late peak accent (called an A accent in the English literature), see Example (79) and (80). Non-final accents can be produced with a late peak accent as well or other pitch accents like rising accents or rise-fall-rise accents. In this study it has been suggested that at least for English, a rising accent (i.e. early Rise B accent or Rise B accent) might simply indicate that it is a non-final accent (and another accent is following).

(79) *What about Fred? What did he eat?* (Second DP focus)

Fred	ate	the [beans]$_F$	
preNuc		Nuc	(order)
Accent		Accent	(accentuation)
A/B		A	(pitch)

(80) *What about Fred? What did he eat?* (First DP focus)

the [beans]$_F$	Fred	ate	
Nuc	postNuc		(order)
Accent	Deaccented		(accentuation)
A			(pitch)

Interestingly, even though contexts like (80) are mainly produced with just one accented syllable, when it is produced with two accented syllables the sequence of pitch accents is similar to any other cases of double accented sentences (i.e. in the double focus case and the second DP focus case), see Example (81). In addition, comparing Example (79) with Example (81) provides us with even more compelling evidence against the idea that pitch accents convey specific meanings: If one were to claim that B accents mark CTs and A accents foci (as Jackendoff (1972) and Büring (1997) do), then one would not expect to find an Example like (81) in which the B accent is placed on a focused constituent rather than on the constituent that contextually serve as the contrastive topic.

(81) *What about Fred? What did he eat?* (First DP focus)

the [beans]$_F$	Fred	ate	
Nuc	postNuc		(order)
Accent	Accent		(accentuation)
A/B	A		(pitch)

From the English data it seems that there is no specific pre-nuclear pitch accent that has to go with a certain manipulation rather the type of pre-nuclear accent is optionally determined by the speaker. The presence or absence of pre-nuclear accents appears to be optional. In our English data the choice of pre-nuclear pitch accent (i.e. A or B accent) is completely arbitrary. One explanation for this arbitrariness could be that the type of pitch accent in pre-nuclear position is indeed completely arbitrary, i.e. there is no pattern to be found. Another more plausible explanation is that the manipulations simply do not capture the phenomenon linked to the pre-nuclear accent. The results from Chapter 3 tell us that the type of pre-nuclear accent in English is not completely arbitrary and some pre-nuclear accents are more prevalent in certain contexts than in others: For instance, the preaccentual Rise-Fall-Rise accent as well as the Rise accent was produced more frequently in contexts with explicit alternative propositions than in other alternative contexts. Moreover, the data from German and Dutch also show that the type of pre-nuclear accent is not completely arbitrary. Contrary to Büring (1997)'s claims, pre-nuclear accents did not consistently mark contrastive topics, in fact rising accents were also produced on constituents that would be considered the Focus of the sentence (Chapter 5). Nonetheless, from the German and Dutch data it became clear that there is a trend for pre-nuclear rising accents as well as nuclear early peak accents to be less appropriate in contexts with no alternative propositions (Chapter 7).

8.3 Prenuclear Accents and Communicative Relevance

In this section I would like to discuss the communicative relevance of pre-nuclear accents and whether pre-nuclear accents add any meaningful information to the discourse or whether they are produced completely optionally. Basically, there are two streams of

thought on the one hand there is the *nuclear-only* view and on the other hand there is the *pre-nuclear-matters* view (Ladd 2008). Briefly, according to the *nuclear-only* view pre-nuclear accents are not relevant in expressing any speaker-intended meaning, whereas the other approach holds the view that pre-nuclear accents do matter and contribute to a meaningful distinction.

Researchers who follow the *nuclear-only view*, have described pre-nuclear accents in English as optional and variable in production (Chodroff and Cole 2018), and such accents do not contribute to any information-structural distinctions Gussenhoven (2011; 2015). Some regard pre-nuclear accents as "ornamental" (Büring 2007)[4] while others have argued that pre-nuclear accents are only used for rhythmic purposesCalhoun (2010). Prenuclear accents are also perceived to be be acoustically less prominent than nuclear accents by listeners according to a study by Cole et al. (2019).

On the other hand there are researchers who argue that pre-nuclear accents are semantically relevant and can be associated with different interpretations of referential and speech act meanings (e.g. Braun and Asano (2013); Petrone and D'Imperio (2011); Petrone and Niebuhr (2014)). In a gating experiment Petrone and Niebuhr (2014) showed that German listeners use the shape, slope and alignment of pre-nuclear pitch accents to distinguish statements from questions. Similar to perception findings, production studies have also found evidence indicating that pre-nuclear accents contribute to information structural distinctions (e.g. Braun and Asano (2013); Féry and Kügler (2008)). Féry and Kügler (2008) for instance investigated the scaling of pitch accents in sequences of accented and deaccented words in relationship to their information status and their place in a tone sequence. Féry and Kügler (2008) found that the peaks of pre-nuclear accents for given arguments (82b) were significantly lower than the peaks of pre-nuclear accents in an all-new context (82a). The conclusion that Féry and Kügler (2008) draw from this result is that pre-nuclear accents can cue givenness and therefore more broadly speaking

[4]It should be noted that this only concerns non-F marked constituents, as CT/F-marked constituents receive a semantically relevant pitch accent

pre-nuclear accents do contribute to meaningful distinctions. Moreover, Féry and Kügler (2008) found the same asymmetry discussed in this study, namely that while constituents in pre-nuclear position can receive a pitch accent (82b), in post-nuclear position such constituents tend to be deaccented (82c).

(82) a. (Féry and Kügler 2008: p. 686) (German, all-new condition)
(Weil der RAMMler dem REIher den HUMmer vorgestellt hat)$_F$
"Because the buck introduced the lobster to the heron."

b. (Féry and Kügler 2008: p. 688) (German, narrow focus in final position)
Weil der Rammler dem Reiher den Hummer (VORgestellt)$_F$ hat.

c. (Féry and Kügler 2008: p. 688) (German, narrow focus in initial position)
Weil der (RAMMler)$_F$ dem Reiher den Hummer vorgestellt hat.

Although it might seem as if these two opposing views are incompatible with each other, a more nuanced image is sketched by Roettger et al. (2021) who provide experimental evidence suggesting that when listeners have access to the full intonation contour they sometimes use pre-nuclear information to anticipate upcoming referential information and sometimes they simply do not. In a number of mouse tracking experiments on German and American English Roettger et al. (2021) tested whether listeners use early intonational information to anticipate upcoming referents. Listeners had to select a speaker-intended referent with their mouse guided by intonational cues, allowing them to anticipate their decision by moving their hand toward the referent prior to lexical disambiguation. While German listeners seemed to ignore early pitch cues, American English listeners were in principle able to use these early pitch cues to anticipate upcoming referents. However, many listeners showed no indication of doing so. These results suggest that there are important positional asymmetries in the way intonational information is integrated, with early information being paid less attention to than later cues in the utterance.

Roettger et al. (2021) argues that the predictive exploitation of pre-nuclear pitch accent is to some extent dependent on the pitch accent type and/or its perceptual salience. This is in line with recent experimental evidence by Braun and Biezma (2019) who show that prominent pre-nuclear accents activate semantic alternatives, while other, less salient, pitch accents do not. In this study we have seen similar results for German and Dutch where both pre-nuclear rising accents and nuclear early peak accents were more acceptable in contexts with alternative propositions than in contexts without alternative propositions.

Just like Roettger et al. (2021), the results from the current study seem to sketch a more nuanced view on the semantic contribution of pre-nuclear accents. While pre-nuclear accents did not seem to play any role in contrastive topic constructions in English (Chapter 2), they did play a role in German (and to a lesser extent in Dutch) (Chapter 5). Based on this observation one might be tempted to conclude that for some languages pre-nuclear accents contribute to a meaningful distinction whereas for other languages they don't. However, as it turned out the manipulations in Chapter 2 simply did not capture the meaning conveyed by the pre-nuclear accent and more pre-nuclear RFR accents were indeed consistently produced in contexts with explicit alternative propositions than in other alternative contexts (see Chapter 3). Although it should be noted that while correlations between pre-nuclear accents and certain semantic manipulations were found, there was a lot of optionality as well. All in all one can conclude that pre-nuclear accents can cue certain meaningful distinctions, but it remains highly optional. This is the same conclusion drawn by Roettger et al. (2021).

8.4 Conclusion

To conclude this dissertation, there are two main points this study makes. First of all, this study does not support a compositional approach to intonational meaning. Throughout

the different experiments we have seen that contrary to claims made by Jackendoff (1972) and Büring (1997), pitch accents do not convey specific meanings. The AB-BA patterns do not convey whether the B-accented or the A-accented functions as a contrastive topic in the utterance, it simply conveys that two constituents evoke alternatives, and non-final ones can be realized with a rise, thus confirming a claim made by Wagner (2012). For German and Dutch, there is a related insight, namely the hat contour does not convey which constituent is a contrastive topic or a focus, it can be used just to convey that two constituents evoke alternatives. With respect to the particular meanings of the AB pattern and the hat pattern, the results were not in line with any of the prior accounts, in particular the claim in Wagner (2012) that AB conveys that some alternative(s) are possibly true and the hat contour conveys that some alternative is true were not confirmed. It should be noted however that Wagner (2012) actually argues that the AB pattern conveys that an alternative speech act is possibly true rather than an alternative propositions. This way of looking at it might help explain why it seems that speakers use the AB pattern when they do (i.e. in for instance incomplete answers but not the contexts in this study), they use the AB pattern when they want to imply another speech act and when they want the interlocutor to draw an additional inference.

Secondly there are also insights about a prosodic asymmetry: a DP that is not used to evoke alternatives is much more likely to be deaccented when it follows a DP that evokes alternatives (i.e. post-nuclear) than when it precedes a DP that evokes alternatives (i.e. pre-nuclear), thus confirming a well known asymmetry in the realization of metrical structure between postnuclear and pre-nuclear expressions (Büring 2016; Ladd 2008; Truckenbrodt 1995; Wagner 2005)

All in all, in this study an attempt was made to get a better understanding of how intonation and meaning relate to each other in AB-BA constructions in English and hat patterns in German and Dutch. Although we seem to have gained a better understanding

of intonation, we have barely scraped the surface and there is still plenty more research to be done.

Bibliography

Abercrombie, D. (1976). Stress and some other terms. *Work in Progress*, 9:51–53.

Adriaens, L. M. H. (1991). *Ein Modell deutscher Intonation: eine experimentell-phonetische Untersuchung nach den perzeptiv relevanten Grundfrequenzänderungen in vorgelesenen Text*. PhD thesis, Technische Universiteit Eindhoven.

Bates, D., Maechler, M., Bolker, B., Walker, S., et al. (2014). lme4: Linear mixed-effects models using eigen and s4. *R package version*, 1(7):1–23.

Bing, J. M. (1979). *Aspects of English prosody*. University of Massachusetts Amherst.

Bishop, J. B. (2013). *Prenuclear accentuation in English: Phonetics, phonology, and information structure*. PhD thesis, UCLA.

Bloomfield, L. (1994). *Language*. Motilal Banarsidass Publ.

Bolinger, D. and Bolinger, D. L. M. (1986). *Intonation and its parts: Melody in spoken English*. Stanford University Press.

Bolinger, D. L. (1958). A theory of pitch accent in english. *Word*, 14(2-3):109–149.

Braun, B. and Asano, Y. (2013). Double contrast is signalled by prenuclear and nuclear accent types alone, not by f0-plateaux. In *INTERSPEECH 2013*, pages 263–266.

Braun, B. and Biezma, M. (2019). Prenuclear l+ h activates alternatives for the accented word. *Frontiers in psychology*, page 1993.

Braun, B. and Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*, 25(7-9):1024–1043.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. PhD thesis, Lund University.

Büring, D. (1997). *The 59th Street Bridge accent: On the meaning of topic and focus*. Routledge.

Büring, D. (1999). Topic. In Bosch, P. and van der Sandt, R., editors, *Focus — Linguistic, Cognitive, and Computational Perspectives*, pages 142–165. Cambridge University Press.

Büring, D. (2007). Semantics, intonation and information structure. *The Oxford handbook of linguistic interfaces*, 445474.

Büring, D. (2016). *Intonation and meaning*. Oxford University Press.

Calhoun, S. (2007). *Information structure and the prosodic structure of English: A probabilistic relationship*. PhD thesis, University of Edinburgh.

Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language*, pages 1–42.

Calhoun, S. (2012). The theme/rheme distinction: Accent type or relative prominence? *Journal of Phonetics*, 40(2):329–349.

Caspers, J. (1997). Testing the meaning of four dutch pitch accent types. In *Fifth European Conference on Speech Communication and Technology*.

Caspers, J. (1999). An experimental investigation of meaning differences between the'early'and the'late'accent-lending fall in dutch. *Linguistics in the Netherlands*, 16(1):27–39.

Caspers, J. (2000). Experiments on the meaning of four types of single-accent intonation patterns in dutch. *Language and speech*, 43(2):127–161.

Caspers, J., van Heuven, V. J., and van Zwol, N. (1998). Experiments on the semantic contrast between the'pointed hat'contour and the accent-lending fall in dutch. *Linguistics in the Netherlands*, 15(1):65–79.

Chodroff, E. R. and Cole, J. (2018). Information structure, affect, and prenuclear prominence in american english. In *Proceedings of INTERSPEECH 2018*, pages 1848–1852. International Speech Communication Association.

Christensen, R. H. B. (2019). ordinal—regression models for ordinal data. R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal.

Cleophas, T. J. and Zwinderman, A. H. (2016). Poisson regression for binary outcomes (52 patients). In *SPSS for Starters and 2nd Levelers*, pages 273–277. Springer.

Cole, J., Hualde, J. I., Smith, C. L., Eager, C., Mahrt, T., and de Souza, R. N. (2019). Sound, structure and meaning: The bases of prominence ratings in english, french and spanish. *Journal of Phonetics*, 75:113–147.

Constant, N. (2012). English rise-fall-rise: a study in the semantics and pragmatics of intonation. *Linguistics and Philosophy*, 35(5):407–442.

Constant, N. (2014). *Contrastive topic: Meanings and Realizations*. PhD thesis, University of Massachusetts Amherst.

Cruttenden, A. et al. (1997). *Intonation*. Cambridge University Press.

Crystal, D. (1969). *Prosodic systems and intonation in English*, volume 1. CUP Archive.

Cutler, A. (1977). The context-dependence of" intonational meanings". In *Thirteenth Regional Meeting, Chicago Linguistic Society*, pages 104–115. CLS.

Cutler, A. (1984). Stress and accent in language understanding and production. *Gibbon, D. and H. Richter (1084)(eds.). Intonation, accent and rhythm. Berlin/New York: Mouton de Gruyter*, pages 77–90.

de Leeuw, J. R. (2015). jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, 47(1):1–12.

De Leeuw, J. R. (2015). jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior research methods*, 47(1):1–12.

Delattre, P. (1966). Les dix intonations de base du français. *French review*, pages 1–14.

Féry, C. (1989). *Prosodic and tonal structure of standard German*. Fachgruppe Sprachwiss. d. Univ.

Féry, C. (1993). The tonal structure of standard german. *Tübingen: Max-Niemeyer*.

Féry, C. (2010). *German intonational patterns*, volume 285. Walter de Gruyter.

Féry, C. (2017). *Intonation and prosodic structure*. Cambridge University Press.

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

Fox, A. et al. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford University Press.

Gartenberg, R. and Panzlaff-Reuter, C. (1991). Production and perception of f0 peak patterns in german. *Studies in German Intonation*, 25:29–113.

Gast, V. (2010). Contrastive topics and distributed foci as instances of sub-informativity. *Comparative and Contrastive Studies of Information Structure (ed. C. Breul and E. Göbbel)*, pages 15–50.

Goodhue, D., Harrison, L., Su, Y. C., and Wagner, M. (2016). Toward a bestiary of english intonational contours. *The Proceedings of the North East Linguistics Society (NELS)*, 46:311–320.

Grabe, E., Gussenhoven, C., Marsi, J. H. E., and Post, B. J. (1998). Preaccentual pitch and speaker attitude in dutch. *Language and speech*, 41(1):63–85.

Grice, M. (1995). Leading tones and downstep in english. *Phonology*, pages 183–233.

Grice, M., Baumann, S., and Benzmüller, R. (2006). Autosegmental-metrical phonology. *Prosodic typology: The phonology of intonation and phrasing*, 1:55.

Grice, M., Baumann, S., and Benzmuller, R. (2007). German intonation in autosegmental-metrical phonology. *Prosodic Typology*, page 55.

Grice, M., Reyelt, M., Benzmuller, R., Mayer, J., and Batliner, A. (1996). Consistency in transcription and labelling of german intonation with gtobi. In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96*, volume 3, pages 1716–1719. IEEE.

Gubian, M. (2013). Functional data analysis for speech research. *(online) lands.let.ru.nl*.

Gundel, J. K. and Fretheim, T. (2004). Topic and focus. *The handbook of pragmatics*, 175(196):12.

Gunter, R. (1972). Intonation and relevance. *Intonation: Selected Readings. D. Bolinger (ed.). Harmondsworth: Penguin*, pages 194–215.

Gussenhoven, C. (1983). Focus, mode and the nucleus. *Journal of linguistics*, 19(2):377–417.

Gussenhoven, C. (1991). Tone segments in the intonation of dutch. In *The Berkeley conference on Dutch linguistics*, volume 1989, pages 139–155. University Press of America, Lanham, MD.

Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology. In *Speech Prosody 2002, International Conference*.

Gussenhoven, C. (2011). Sentential prominence in english. *The Blackwell companion to phonology*, pages 1–29.

Gussenhoven, C. (2015). Does phonological prominence exist? *Lingue e linguaggio*, 14(1):7–24.

Gussenhoven, C. (2016a). Analysis of intonation: The case of mae_tobi. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1).

Gussenhoven, C. (2016b). Foundations of intonational meaning: Anatomical and physiological factors. *Topics in Cognitive Science*, 8(2):425–434.

Gussenhoven, C. et al. (2005). Transcription of dutch intonation. *Prosodic typology: The phonology of intonation and phrasing*, 118:145.

Halliday, M. A. K. (1967). *Intonation and grammar in British English*. De Gruyter Mouton.

Hamblin, C. (1973). Questions in montague english. *Foundations of Language*, 10(1):41–53.

Höhle, T. N. (1991). On reconstruction and coordination. *Beiträge zur deutschen Grammatik*, page 311.

Jackendoff, R. S. (1972). Semantic interpretation in generative grammar. *ERIC*.

Jackman, S., Tahk, A., Zeileis, A., Maimone, C., Fearon, J., Meers, Z., Jackman, M. S., and Imports, M. (2015). Package 'pscl'. *Political Science Computational Laboratory*, 18(04.2017).

Jacobs, J. (1982). *Syntax und Semantik der Negation im Deutschen*, volume 1. W. Fink.

Jassem, W. and Gibbon, D. (1980). Re-defining english accent and stress. *Journal of the International Phonetic Association*, 10(1-2):2–16.

Jun, S.-A. (2007). *Prosodic typology: The phonology of intonation and phrasing*, volume 1. Oxford University Press on Demand.

Keijsper, C. E. et al. (1984). Vorm en betekenis in nederlandse toonhoogtecontouren. In *Forum der Letteren*, volume 25, page 2.

Kingdon, R. (1958). *The groundwork of English intonation*. Longmans.

Kohler, K. J. (1990). Macro and micro f0 in the synthesis of intonation. *Papers in Laboratory Phonology I*, pages 115–138.

Kohler, K. J. (1991). Terminal intonation patterns in single-accent utterances of german: phonetics, phonology and semantics. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)*, 25(1):15–185.

Krifka, M. (1998). Additive particles under stress. In *Semantics and Linguistic Theory*, volume 8, pages 111–128.

Ladd, D. R. (2008). *Intonational Phonology*. Cambridge University Press.

Ladd, D. R., Schepman, A., White, L., Quarmby, L. M., and Stackhouse, R. (2009). Structural and dialectal effects on pitch peak alignment in two varieties of british english. *Journal of Phonetics*, 37(2):145–161.

Ladd Jr, D. R. (1980). *The structure of intonational meaning: Evidence from English*. Indiana University Press.

Laver, J. and John, L. (1994). *Principles of phonetics*. Cambridge university press.

Lee, C. (2008). Contrastive (predicate) topic, intonation, and scalar meanings. In *Topic and focus*, pages 151–175. Springer.

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

Liberman, M. and Sag, I. (1974). Prosodic form and discourse function. In *Chicago Linguistics Society*, volume 10, pages 416–427.

Liberman, M. Y. (1978). *The intonational system of English*. PhD thesis, Massachusetts Inst. of Technology Cambridge.

Ludwig, R. A. (2006). *Information structure and scope inversion*. PhD thesis, Diplomarbeit. Universität Potsdam.

Moulton, W. (1982). Sentence perspective: linear order, stress, and intonation. *Ms. Revision of paper presented at the Penn Linguistics Club*, 10.

Newman, S. S. (1946). On the stress system of english. *Word*, 2(3):171–187.

Niebuhr, O. (2003). Perceptual study of timing variables in f0 peaks. In *Proc. 15th ICPhS, Barcelona*, pages 1225–1228.

O'Connor, J. D. and Arnold, G. F. (1973). *Intonation of colloquial English*. Longman.

Ohl, C. K. and Pfitzinger, H. R. (2009). Compression and truncation revisited. In *Tenth Annual Conference of the International Speech Communication Association*.

Oshima, D. (2005). Morphological vs. phonological contrastive topic marking. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, volume 41, pages 371–384. Chicago Linguistic Society.

Petrone, C. and D'Imperio, M. (2011). From tones to tunes: Effects of the f 0 prenuclear region in the perception of neapolitan statements and questions. In *Prosodic categories: Production, perception and comprehension*, pages 207–230. Springer.

Petrone, C. and Niebuhr, O. (2014). On the intonation of german intonation questions: The role of the prenuclear region. *Language and Speech*, 57(1):108–146.

Pierrehumbert, J. and Hirschberg, J. B. (1990). The meaning of intonational contours in the interpretation of discourse. *Intentions in Communication*, pages 271–311.

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

Pike, K. L. (1945). *The Intonation of American English.* Ann Arbor : University of Michigan Press.

Pitrelli, J. F., Beckman, M. E., and Hirschberg, J. (1994). Evaluation of prosodic transcription labeling reliability in the tobi framework. In *Icslp*, volume 1, pages 123–6.

Ramsay, J., Hookers, G., and Graves, S. (2009). *Functional data analysis with R and MATLAB.* Springer Verlag, New York, NY.

Rathcke, T. and Harrington, J. (2006). Is there a distinction between h+! h* and h+ l* in standard german? evidence from an acoustic and auditory analysis. In *Proceedings of Speech Prosody*.

Rathcke, T. and Harrington, J. (2007). The phonetics and phonology of high and low tones in two falling f0-contours in standard german. In *Eighth Annual Conference of the International Speech Communication Association*.

Rathcke, T. and Harrington, J. (2010). The variability of early accent peaks in standard german. *Laboratory Phonology*, 10:533–555.

Rietveld, T. and Gussenhoven, C. (1995). Aligning pitch targets in speech synthesis: effects of syllable structure. *Journal of Phonetics*, 23(4):375–385.

Roettger, T. B., Franke, M., and Cole, J. (2021). Positional biases in predictive processing of intonation. *Language, Cognition and Neuroscience*, 36(3):342–370.

Rooth, M. (1985). *Association with focus*. PhD thesis, Dept. of Linguistics, University of Massachusetts, Amherst.

Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics*, 1(1):75–116.

Schubiger, M. (1958). *English intonation, its form and function*. M. Niemeyer Verlag.

Selkirk, E. O. (1986). *Phonology and syntax: the relationship between sound and structure*. MIT press.

Silverman, K. E., Beckman, M. E., Pitrelli, J. F., Ostendorf, M., Wightman, C. W., Price, P., Pierrehumbert, J. B., and Hirschberg, J. (1992). Tobi: A standard for labeling english prosody. In *ICSLP*, volume 2, pages 867–870.

Sluijter, A. (1995). Phonetic correlates of stress and accent. *PhD Thesis, Holland Institute of Generative Linguistics*.

Sluijter, A. M. and Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical society of America*, 100(4):2471–2485.

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

Sweet, H. (1929). *The sounds of English: an introduction to phonetics*. Clarendon Press.

Syrdal, A. K. and McGory, J. (2000). Inter-transcriber reliability of tobi prosodic labeling. In *Sixth International Conference on Spoken Language Processing*.

't Hart, J. and Collier, R. (1975). Integrating different levels of intonation analysis. *Journal of Phonetics*, 3(4):235–255.

't Hart, J., Collier, R., and Cohen, A. (1990). *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge University Press.

Team, R. C. (2008). R: A language and environment for statistical computing. *Journal of Statistical Software*, 25(1).

Torreira, F. and Valtersson, E. (2015). Phonetic and visual cues to questionhood in french conversation. *Phonetica*, 72(1):20–42.

Truckenbrodt, H. (1995). *Phonological phrases–their relation to syntax, focus, and prominance*. PhD thesis, Massachusetts Institute of Technology.

Van der Hulst, H. (2011). Pitch accent systems. *The Blackwell companion to phonology*, 2:1003–1026.

Van Hoof, H. (2003). The rise in the rise-fall contour: does it evoke a contrastive topic or a contrastive focus? *Linguistics*, 41(3; ISSU 385):515–564.

Venditti, J. J. (1997). Japanese tobi labelling guidelines. *Papers from the Linguistics Laboratory, Ohio State University Working Papers in Linguistics*, 50:127–162.

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

Wagner, M. (2008). A compositional analysis of contrastive topics. In *Proceedings of NELS*, volume 38, pages 415–428.

Wagner, M. (2012). Contrastive topics decomposed. *Semantics and Pragmatics*, 5:8–1.

Wagner, M., McClay, E., and Mak, L. (2013). Incomplete answers and the rise-fall-rise contour. In *Workshop on the Semantics and Pragmatics of Dialogue (SemDial)*, volume 17, pages 140–149. Citeseer.

Ward, G. and Hirschberg, J. (1985). Implicating uncertainty: The pragmatics of fall-rise intonation. *Language*, pages 747–776.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, 91(3):1707–1717.

Zou, G. (2004). A modified poisson regression approach to prospective studies with binary data. *American journal of epidemiology*, 159(7):702–706.