

# German as a tone language<sup>1</sup>

Caroline Féry & Frank Kügler, University of Potsdam

## 1. Introduction

The scaling of high tones in the syntactic domain called "middle field" of German sentences is examined experimentally when the number of verbal arguments, their focus structure and the word order is varied. It is shown that variations in pitch height are influenced by three kinds of effect. The first one is well-known: information structure, like givenness, wide focus (all-new sentences) or narrow focus, may raise an accent, lower it or even deaccent constituents. The second kind of effect is more surprising: tonal scaling is influenced by the interaction of tones with each other. The most well-known interaction is probably downstep, a significantly lower scaling of high tones as compared to declination. But we show that other types of tonal interactions have a considerable influence on the scaling of individual tones, as well. Two dissimilatory effects, H-tone raising and H-lowering, have been described in a number of tone languages (for H-raising, see Laniran & Clements 2003 for Yoruba and Xu 1997 for Mandarin Chinese), but never to our knowledge, for an intonation language like German. And finally, the number of pitch accents to be realized in a sentence may also affect the height of an initial high tone. This third effect, called preplanning or long-distance anticipation, has been described for tone languages (Stewart 1993 for Dschang and Ebrié, Rialland & Somé 2002 for Dagara, to cite just a few), and to a lesser extent for intonation languages (see Thorsen 1983 for Danish, and the rejection of preplanning by Pierrehumbert 1980; recently for Romance languages, Prieto et al. 2006). In short, this paper demonstrates that, beside well-known effects on tone scaling coming from information structure, strictly tonal effects, which so far have been described only for tone languages, exist in German, as well. We integrate our experimental findings in a new model of register and pitch accent scaling involving two different changes in tone height. Information structure manipulates the height of reference lines (Bruce 1977, van den Berg et al 1992, Féry & Truckenbrodt 2005), which define register domains associated with prosodic domains. H-raising and H-lowering changes the value of the individual tones inside these registers.

The sections of this paper are organized as follows. Section 2 provides a theoretical background for the issues addressed in the paper. The experimental make-up is described in section 3. Sections 4 and 5 present the results of the experiment and a phonetic analysis of the findings. Section 4 first gives an overview of the tonal patterns of the elicited sentences as influenced by the information structure of the sentence. Section 5 discusses the same tonal patterns, but this time taking the tonal effects, especially H-raising and H-lowering influences into account. Section 6 presents our model of register and pitch accent scaling taking our results into considerations. Section 7 concludes the paper.

---

<sup>1</sup> This paper is part of the project A1 of the SFB632 on Information Structure, and of the project Prosody in Parsing, both financed at the University of Potsdam by the Deutsche Forschungsgemeinschaft. Thanks to Gisbert Fanselow, Robin Hörnig, Shin Ishihara, Anja Mietz, Gerrit Otten, Lisa Selkirk, Hubert Truckenbrodt, Thomas Weskott and Sabine Zerbian for helpful discussions and comments. Usual disclaimers apply. Our deep gratitude goes to Anja Arnhold, Susanne Genzel, Esther Sommerfeld for technical support.

## 2. Background

Researchers agree to classify German as an intonation language (see the classifications in Ladd 1996, Gussenhoven 2004 and Jun 2005). It has lexical accents and uses different types of post-lexical pitch accents and pitch accent sequences, as well as boundary tones for the expression of pragmatic contrasts (von Essen 1956, Pheby 1980, Uhmann 1991, Féry 1993, Mayer 1997, Peters 2005). Different aspects of German dialects are well studied (Peters 2004, Gilles 2005, Kügler 2005, Barker 2002), but here, we restrict ourselves to Standard German as spoken in the Berlin-Brandenburg region. Until now, the bulk of the analysis of the tonal structure of this language has been laid on the tone inventory, the alignment of these tones with segments, as well as variation in the use of tones. Tone scaling across accents and across phrases has not been extensively studied until now, though some research has been done on the subject (see Selkirk 2006, Katz & Selkirk 2005, Truckenbrodt 2002, Féry & Truckenbrodt 2005, Kügler, Féry & van de Vijver 2003). A number of studies on German have specifically investigated downstep (Grabe 1998, Truckenbrodt 2004). To our knowledge, however, no study has been published so far that addresses tone scaling in relationship with information structure more than in passing, and purely tonal effects like H-raising and H-lowering have been left unnoticed.

In declarative sentences like the ones used in the experiment, pitch accents are typically realized in two ways which differ sharply with each other: prenuclear accents have a rising contour, transcribed in (1) as a bitonal L\*H tone sequence and nuclear accents are falling and are transcribed as H\*L (Uhmann 1991, Féry 1993, Grabe 1998, Kügler 2005, Peters 2005) or HL\* in some analyzes (Grice et al 2005, Truckenbrodt 2004). This kind of transcription is coached in an autosegmental-metrical analysis, originally proposed by Pierrehumbert (1980) for English, and applied to German by the authors mentioned above (see in particular the most recent proposal of GToBI, Grice et al. 2005, and references therein).

- L\*H                    L\*H                    H\*L                    L<sub>i</sub>
- (1) Eine FRAU hat einem KIND ein BUCH gegeben.  
A woman has a.DAT child a.ACC book given.

It is a matter of debate how to express the low contour on the verb *gegeben* in (1), or more generally the low and flat postnuclear stretch in a declarative sentence. The fall in such a sentence happens immediately after the last accent. Two main options compete with each other to account for this contour.<sup>2</sup> In the first one, the L portion of the last bitonal tone reaches the baseline. Since no further tone follows, the pitch remains low until the end of the sentence. The second option is to associate a low boundary tone (L<sub>i</sub>) with the end of the sentence and to align it with the postnuclear syllable. The low boundary tone extends from the end of the sentence backward to the nuclear accent. This is in line with recent proposals by Gussenhoven (2004) and his distinction between association and alignment of tones with syllables.

Pitch accents are associated with prominent positions in the sentence. Default sentence accent assignment, as well as accents motivated by information structure have been examined extensively for German (for instance by Bierwisch 1968, Fuchs 1976, Gussenhoven 1992, Jacobs 1988, von Stechow & Uhmann 1986, Cinque 1993, Buring 2001, and Féry & Samek-Lodovici 2006). The sentences investigated in the present

---

<sup>2</sup> A solution in the line of Grice, Ladd & Arvaniti (2000) would involve the existence of phrase accents. But in our data nothing bear on this issue.

paper consist of one to three verbal arguments followed by a final participle and auxiliary (see the examples in (3) below). In order to capture accent assignment in this kind of sentences, the following principles are formulated. In an all-new sentence, accents are assigned on the basis of the syntactic structure of the sentence. Every argument of a verb is accented and the verb itself may be accented or not, depending on the phrasal integration of the verb and its immediately preceding argument, an optional process (see for instance Fuchs 1976, Jacobs 1993 and Gussenhoven 1992). In sentences with narrow foci, main accent is assigned to the last focused constituent, in our sentences, any argument or the verb. Prenuclear arguments are always accented and postnuclear words are always deaccented.

Information structure is an important component of our study. The notions used here are restricted to focus and givenness, which are regulated by the context in which the experimental sentences were embedded. Focus appears in two variants: wide and narrow (Ladd 1980). In a wide or whole focused sentence, all elements are new and have not been mentioned in the preceding context. A narrow focus is induced by a context asking explicitly for one or more arguments, or for the verb. A semantically informed definition of focus and givenness is beyond the scope of this paper (see Rooth 1985, 1992 and Schwarzschild 1999 among others), but since we rely on an impoverished conceptualization, the definitions in (2) can serve as an orientation.

(2) a. Focus: the information explicitly asked for by the preceding question is focused. It can be the whole sentence (wide focus, which we refer to as the "all-new sentence condition" henceforth) or just some part of the sentence (narrow focus).

b. Givenness: The information already mentioned in the preceding question is given, and only repeated in the target sentence.

Turning now to the tonal phenomena of relevance for this paper, it must be observed that pitch declines over the time course of an utterance. This downward trend has been studied under two headings: downstep and declination. Declination has its physiological basis in a gradually decrease of subglottal air pressure (see for instance Lieberman 1967, Cohen, 't Hart 1967, Cooper & Sorensen 1981, Grønnum 1992). Declination is a temporally bounded phenomenon. It is not considered any further in this paper.

Phrase-internally, the decline in pitch affects both the lower and the upper part of the pitch register. Whereas the lower part of the register (the bottom line) is relatively stable, downstep affects the upper part of the pitch register (the top line) where tonal targets may be realized significantly lower or higher than in case of neutral declination (for downstep, see among others Pierrehumbert 1980, Liberman & Pierrehumbert 1984, Ladd 1984, 1996, as well as Maeda, 1976, Thorsen, 1983). According to Pierrehumbert (1980) and Liberman & Pierrehumbert (1984), scaling of tones in English takes place from left to right, downstepping a high tone relatively to a preceding high tone. The calculation of the phonetic height of a tone takes the local left and right phonological context into account, as well as the left phonetic context. According to German intonation, we expect downstep to be the neutral realisation of an all-new sentence.

The reverse effect, upstep, has also been claimed to exist in German. Truckenbrodt (2002) proposes that upstep is an optional phenomenon limited to the last pitch accent in a non-final intonation phrase. He analyzes upstep as an undoing of a downstepped sequence. In Truckenbrodt's account, upstep must be preceded by downstepped pitch accents and followed by an intonation phrase (see also Féry & Truckenbrodt 2005). Since the study presented here concentrates on sentences consisting of only one intonation phrase, we never find this kind of upstep in our data, although we do find a similar effect of raising of the last accent in an intonation phrase.

Considering the effects of information structure on the realisation of accents, two different phonetic implementations are known. While focus generally boosts accents in a number of languages (Bartels & Kingston 1994, Brown et al. 1980, Ladd 1993, Féry & Ishihara 2005, Ishihara & Féry 2006), given constituents in prenuclear position are realized lower (Baumann 2006) or are deaccented in postnuclear position (Ladd 1980). These effects are shown in detail for German in section 4.1 below. Our analyses confirm these general findings.

Two further tonal effects of relevance for the following are H-raising and H-lowering. H-raising is an effect which has been until now identified only in tone languages. It is analyzed by Xu (1997) for Mandarin Chinese as an anticipatory dissimilatory effect. A high tone is realized even higher as anticipation of a following low tone. Laniran & Clements' (2003) study of tonal interactions in Yoruba has significantly influenced the present paper. Laniran & Clements report the results of production experiments in Yoruba and give a careful analysis of the speech of four speakers. They propose that the overall shape of an  $F_0$  contour is the result of a compromise between different tendencies in the tonal pattern. They find two effects driving H tones to progressively lower their pitch, downstep, and downdrift, a tone-specific declination. But high tones are sometimes higher than predicted by the two lowering effects. First, H-raising happens before L, and second, register adjustment may take place, like reset of high tones in a long sequence of high tones or raising an initial H as an anticipation of a long high tone sequence. Speakers use the same strategies in different quantities, conspiring to prevent downstepped high tones from penetrating  $F_0$  levels reserved for M and L tones. They give a compositional analysis of tone scaling in which the melodic contour of an utterance is the result of a number of interacting factors. See also Clements (1990), Berg et al. (1992), Ladd (1990) and others for this kind of interacting effect of tones and register on the scaling of individual pitch heights.

### 3. Experiment

#### 3.1 Material

The aim of the experiment reported in this section was to investigate the scaling of pitch accents in sequences of accented and deaccented words in relationship with their information status, their place in a tone sequence and the number of realized accents. The sentences used in the experiment have a simple syntactic pattern: They are verb-final sentences introduced by a complementizer *weil* 'because', which contain one, two or three arguments plus the verb. A single argument is always a nominative. When there are two arguments, they are nominative and accusative, or nominative and dative. Sentences with three arguments contain a nominative, a dative and a accusative. In the experiment, all sentences were introduced by a context which could be just one question (in case of wide focus) or an introductory sentence followed by a question. The arguments were masculine animal names, so that Case was unambiguously recognizable on the article (German feminine and neuter articles have the same morphological form in nominative and accusative). All five animal names were trochaic with a final schwa syllable (*Hammel* 'wether', *Hummer* 'lobster', *Löwe* 'lion', *Rammler* 'buck', *Reiher* 'heron'). Only particle verbs were used, with detachable and stressed particles. The complex verbs were always in the participial form and located at the sentence-final position (*angefangen* 'begun', *eingeladen* 'invited', *nachgelaufen* 'followed', *vorgestellt* 'introduced'). They vary in the number of arguments they require and in the Cases they govern. Examples of the sentences are given in (3). The material marked by curly brackets is a translation of the context sentences. The target sentences are in italic, underlined constituents indicate focused material. N stands for nominative, A for accusative, D for dative and V for verb.

(3) Some examples

a. **NDV** (focus on two arguments but the verb)

{The animals like to play ‘catch’. One animal has to follow another one. Why were they happy this time?}

*Weil der Hammel dem Rammler nachgelaufen ist.*

because the wether the buck followed has

‘Because the wether has followed the buck’

b. **NAV** (all-new condition)

{Why were the animals happy?}

*Weil der Hammel den Rammler eingeladen hat.*

‘Because the wether has invited the buck’

c. **NADV** (focus on the dative complement)

{The wether wanted to present the buck to the lion. Why didn’t he do this?}

*Weil der Hammel den Rammler dem Hummer vorgestellt hat.*

‘Because the wether presented the buck to the lobster.’

d. **NV** (focus on the verb)

{The animals don’t like fight. Why are they angry with the wether?}

*Weil der Hammel angefangen hat.*

‘Because the wether started (a fight).’

Three parameters were systematically varied:

1. Number of arguments: between one and three:

- Nominative-verb (NV)
- Nominative-accusative-verb (NAV) or nominative-dative-verb (NDV)
- Nominative-dative-accusative-verb (NDAV)

2. Word order:

NV, NAV, ANV, NDV, DNV, NDAV, NADV, ADNV, ANDV, DANV, DNAV

3. Given-new status of the arguments and the verb

- the whole sentence is focused (first column in table 1)
- one argument is focused (second and third column in table 1)
- only the verb is focused (third column in table 1)
- the arguments are focused to the exception of the verb (last column in table 1)

Altogether 26 conditions were created, which are listed here:

**Table 1:** Recorded material in overview, focused constituents are underlined<sup>3</sup>

All-new	Narrow focus on one argument	Narrow focus on the verb	Focus on all arguments
<u>NV</u>	<u>NV</u>	<u>NV</u>	
<u>NAV</u>	<u>NAV</u>	<u>NAV</u>	<u>NAV</u>
	<u>N<del>A</del>V</u>		
	<u>A<del>N</del>V</u>		
<u>NDV</u>	<u>NDV</u>	<u>NDV</u>	<u>NDV</u>

<sup>3</sup> The missing constellations imply very unnatural possibilities, like focus on an initial scrambled constituent.

	<u>NDV</u>		
	<u>DNV</u>		
<u>NDAV</u>	<u>NDAV</u>	<u>NDAV</u>	<u>NDAV</u>
	<u>NDAV</u>		
	<u>NADV</u>		
	<u>ADNV</u>		
	<u>ANDV</u>		
	<u>DANV</u>		
	<u>NDAV</u>		
	<u>DNAV</u>		

Each condition was realized in five versions, in which the nouns were permuted in a systematic way, but the verb remained constant. In total, this results in 130 sentences (26 conditions x 5 renditions) per speaker.

### 3.2 Recordings

The entirety of the data used in this paper has been collected in one experiment, run individually with 18 speakers. All were female students at the University of Potsdam, Germany. They were monolingual speakers of German in their twenties, coming from the Northern area of Germany. They were reimbursed for their time.

From the total of 2340 sentences uttered by the 18 speakers (130 each) 63 were not considered in the final results because of immeasurable accents, again mostly due to creaky voice. Altogether 2277 sentences were retained for analysis.

The target sentences were recorded in a sound-proof box on a DAT tape-recorder. A set of instructions familiarized the subjects with the process and made them practice with four examples. After the instructional part, the experimenter left the subject alone in the room. The subject went through the experiment in the form of a Powerpoint presentation in a self-paced manner. The speakers read the sentences on a screen as the answers to questions which were presented both visually and acoustically over headphones: They heard and read a question on a computer screen, pressed the return key, and read aloud a target sentence presented on the next slide. The items to be accented were underlined in order to minimize errors. The context sentences had been recorded preliminarily. They were spoken by the second author, a trained phonetician, native speaker of Standard German in his thirties, also coming from the northern part of Germany. He was recorded in a sound-proof booth on a DAT recorder. He spoke naturally, in a conversational tempo.<sup>4</sup> The target sentences of the present experiment were intermingled with filler sentences, coming from other experiments.

The recorded sentences were digitized at a sampling rate of 16kHz. They were analyzed using the acoustic speech analysis software Praat (Boersma & Weenink 2006). The recordings were partly automatically and partly manually divided into labeled substrings with the help of spectrograms and acoustic inputs. Obvious errors of the F<sub>0</sub> algorithm (for instance octave jumps) have been corrected by hand, and the contour has been smoothed using the Praat smoothing algorithm (frequency band 10 Hz) to diminish microprosodic perturbations. All frequency measurements were semi-automatically done using a script that detects the highest F<sub>0</sub> value within a given domain. The domains for measurements were the complementizer, each argument (article plus noun), the participle and the auxiliary *hat* 'has'. An example of the segmentation appears in (4).

(4) # Weil # der Hammel # den Hummer # eingeladen # hat #

<sup>4</sup> Exemplary sentences can be downloaded from the following website: <http://www.sfb632.uni-potsdam.de/homes/fery/aktuell.html>

because the.NOM wether the.ACC lobster invited has

Technically the analysis was done in a number of steps. In a first step, a Praat script located  $F_0$  maxima in each of the domains. For example, in (4), five  $F_0$  maxima were identified. In a second step, the result of the Praat script was hand-edited to correct spurious labeling. The authors inspected the tone labels against the  $F_0$ -track, the sub-string-divisions, an auditory impression, and the spectrogram. Where the more permissive Praat script had assigned a H label that was not in the position where the more narrow criteria above would place it (because of obvious errors of the algorithm), the label was manually moved. In a third step, another Praat script recovered the  $F_0$ -values at the positions of the tone labels as well as the tone labels themselves, and collected them in a table. Finally, for the duration measurements, a Praat script collected the duration of each domain (cf. (4)).

The tables thus assembled were sorted into utterances of which the values were included in the overall evaluation. And finally, a series of statistical analyses were performed.

## 4. Results 1: Overview of the information structure driven results

The presentation of the results proceeds in a number of steps. In this section, the influence of information structure on pitch and duration is examined. It is shown that narrow focus raises the high part of a pitch accent, and givenness lowers it. As will become clear, such a simple model only accounts for part of the data. In section 5, further effects of tonal interferences are discussed. In the first section, we concentrate on pitch and discuss the realizations of all-new sentences (4.1). As will be shown shortly, this is the most complex case, implying unexpected variation in the way speakers scaled the pitch accents. In a second step (4.2), results for pitch of sentences involving narrow focus are presented. The effect of information structure on duration are summed up in section 4.3. The prefinal subsection shows briefly that word order did not affect pitch height and duration in this experiment (4.4). The results of this section are summed up in section 4.5.

### 4.1 Effect of focus structure

#### 4.1.1 All-new sentences

In the case of wide-focus (conditions NV, NAV, NDV, NDAV in the first column of table 1 above), we expected all arguments to be accented, but in fact, some variation in the assignment of a pitch accent to the verb was observed (cf. section 2). Our second expectation concerned the scaling of these pitch accents: we expected a downstepped pattern in all four conditions. Again our expectations were only partially met.

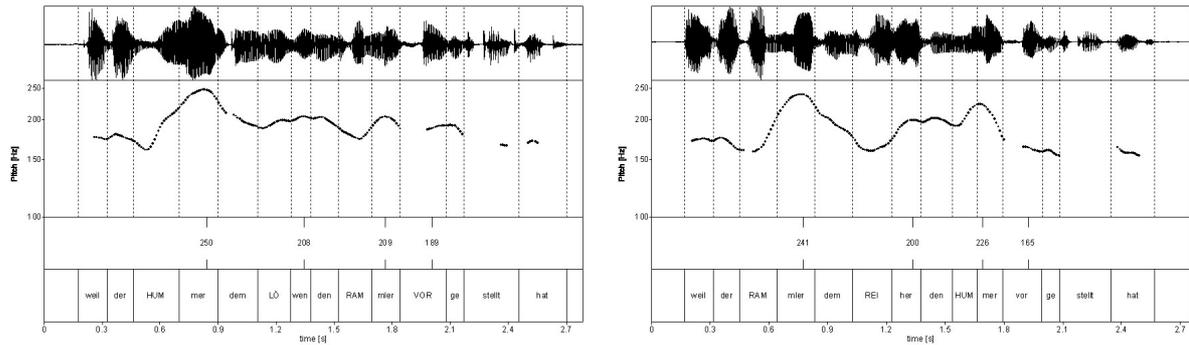
In (5), both versions of a wide focused sentence are illustrated (in one version, the verb is accented and in the other the verb is unaccented). Small capitals stand for a bitonal pitch accent. In the first version (5a), the verb is accented, whereas it is unaccented in (5b). The subscripted 'F' stands for 'focus.' Recall that a sentence which is entirely focused is meant to be uttered as the answer to a question in which neither the participants nor the verb has been mentioned in the preceding question. They differ from the sentences with a narrowly focused constituent in lacking a given (previously mentioned) constituent.

(5) Why were the animals happy?

a. (Weil der HAMMEL den RAMMLER EINGELADEN hat)<sub>F</sub>

- b. (Weil der HAMMEL den RAMMLER eingeladen hat)<sub>F</sub>  
 'Because the wether has invited the buck.'

The variation in the accent pattern is not surprising given that a final verb in German can be 'integrated' into the prosodic domain of a preceding argument (Fuchs 1976, Jacobs 1993, Gussenhoven 1992), but does not need to (see the discussion in section 2). Fig.1a and b show examples of wide-focused sentences, realized by the same speaker but in two different items of the same condition. Concentrating on the verb, it is clearly visible that it has a pitch accent in the first pitch track, but not in the second one.<sup>5</sup>



**Fig.1:** Two versions of all-new sentences with three arguments (NDAV): Fig.1a (left panel) shows an accented verb. Fig.1b (right panel) shows an unaccented verb.

Altogether 348 realizations of wide focused sentences were analyzed in the four all-new sentence conditions (see table 2 for an overview of the number of sentences in each of the four conditions). As explained, two classes of sentences could be distinguished: an accented verb was realized in 71% (247) of the sentences and an unaccented verb in 29% (101).<sup>6</sup> The decision as to which sentences belong to which group was made by the authors, first individually, and then together. In the few unsettled cases resulting from this method, agreement was readily found. The ratio between accented and unaccented verb is illustrated in (6). As can be seen from these data, the verb was accented more often in a short sentence than in a long one.<sup>7</sup>

(6) Distribution of pitch accent on the verb in the all-new sentence conditions.

	Accented verb	Unaccented verb
NDAV ( <i>vorge</i> stellt)	50% (45)	50% (45)
NDV ( <i>nach</i> gelaufen)	75% (64)	25% (21)
NAV ( <i>ein</i> geladen)	73% (61)	27% (22)
NV ( <i>an</i> gefangen)	86% (77)	14% (13)

Turning now to the scaling of the accents, two large classes of cases must be distinguished: In 46% of the sentences, a regular downstep takes place, in which each accent is lower than the preceding one. But in the remaining 54%, the pattern presented raising of an accent. The raised constituent was either the preverbal argument, in which case the verb was unaccented (see Fig.1b), or on the verb, in which case, of course, the

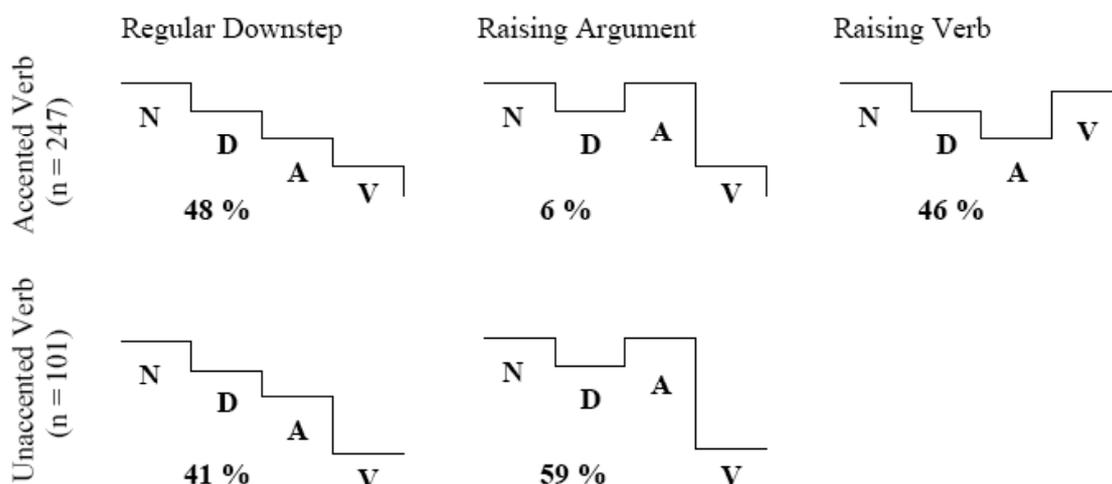
<sup>5</sup> All speakers showed variation in the accenting in the verb, a fact corroborating that the presence vs. absence of an accent is in free variation.

<sup>6</sup> The large number of accented verbs in the wide focus sentences may be an artifact of our experimental design: the focused constituents were underlined. It may be the case that the informants felt compelled to actually accent all underlined words. We suspect that in a natural situation more unaccented verbs would be realized.

<sup>7</sup> In all relevant tables, the conditions NV appears at the bottom. These short sentences often displayed a different pattern than the longer ones.

verb was accented. In short, the last accent interrupts downstep and is realized much higher than predicted by a regularly descending pattern. It will be shown in section 5.2. that this raising is in fact H-raising preceding a final drop in pitch. In a few cases (15 realizations altogether), the raising took place on the final argument preceding an accented verb.

Schematically, all wide-focused realizations are summed up in Fig. 2. From this graph, it can be seen that the variation is limited to the final part of the sentence. The initial arguments are nearly always downstepped.<sup>8</sup>



**Fig.2:** The realization of all-new sentences.

Table 2 classifies the data in finer-grained categories. First, when the verb was accented ( $n = 247$ ), 48% (118) of the sentences presented a regular downstep (ds), and 46% (114 sentences) had a raising on the verb or, less often, on the preverbal argument (6%, 15 sentences).

In the sentences with unaccented verb ( $n = 101$ ), 59% (60 sentences) had a raising on the argument preceding the verb. In 41% of the sentences, a regular downstep could be observed (41 sentences).

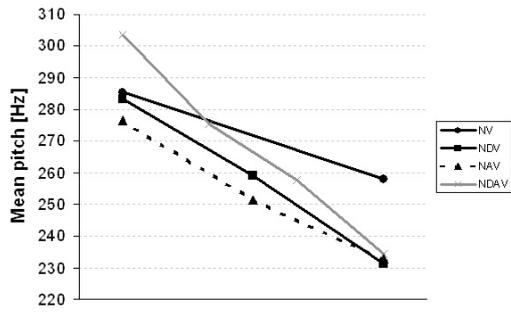
Table 2: Distribution of realizational options of the all-new condition.

	Accented verb (71%)			Subtotal	Unaccented verb (29%)			n
	Regular downstep	Raising argument	Raising verb <sup>9</sup>		Regular downstep	Raising argument	Subtotal	
<u>NAV</u>	28	6	27	61	6	16	22	83
<u>NDV</u>	34	3	27	64	9	12	21	85
<u>NDAV</u>	23	6	16	45	13	32	45	90
<u>NV</u>	33	-	44	77	13	-	13	90
Total	118 (48%)	15 (6%)	114 (46%)	247	41 (41%)	60 (59%)	101	348

Figure 3 shows the average F0 value for each accent, as well as for the unaccented verb. The values in the cells and on the graphs, if nothing indicates the contrary, are averaged high tones for all speakers and all utterances.

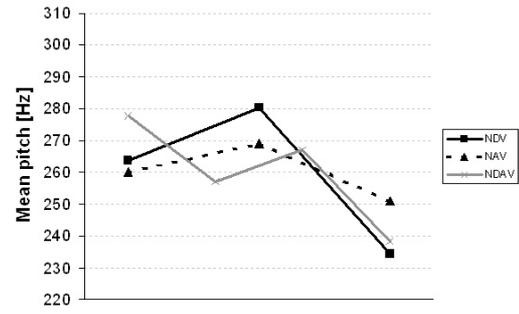
<sup>8</sup> There were a few sentences in the corpus where the arguments were realized either at the same height or where the later arguments were pronounced even higher than the early ones. Their number is insignificant, and we refrain to group them in a separate group. Instead we decided to treat them as lacking downstep, and they appear in the group of sentences with raised arguments.

<sup>9</sup> The raising of the verb is not due to a final rise (list intonation or an indication of continuation). In general, in the test sentences the verb was always followed by an auxiliary, and the auxiliary always showed low pitch.



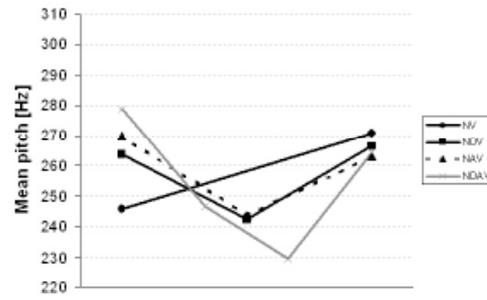
	NDV	NAV	NDAV	NV
Nom	283	276	304	285
Dat	259		275	
Acc		251	258	
Verb	231	233	234	258

(a) regular downstep, accented verb



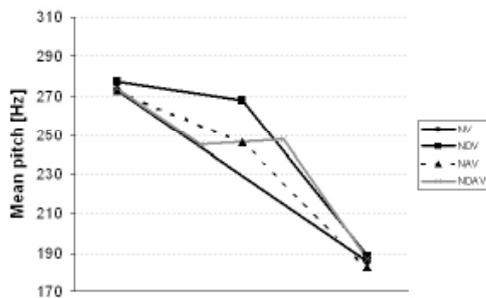
	NDV	NAV	NDAV	NV
Nom	264	260	278	
Dat	280		257	
Acc		269	267	
Verb	234	251	238	

(b) upstep argument, accented verb



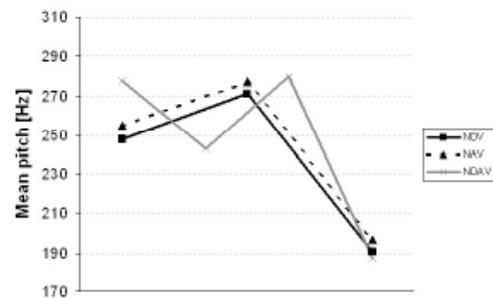
	NDV	NAV	NDAV	NV
Nom	264	270	279	246
Dat	242		247	
Acc		244	230	
Verb	266	263	264	271

(c) upstep verb, accented verb



	NDV	NAV	NDAV	NV
Nom	277	273	274	273
Dat	268		245	
Acc		246	248	
Verb	189	183	187	186

(d) regular downstep, unaccented verb



	NDV	NAV	NDAV	NV
Nom	248	255	278	
Dat	271		243	
Acc		277	280	
Verb	191	197	187	

(e) upstep argument, unaccented verb

**Fig. 3:** Representation of pitch peak realizations of all-new sentences.

Table 3 and 4 sum up average pitch values for all speakers and all items in the four conditions, separated in two groups: verb accented and verb unaccented. A1 stands for the initial argument, A2 and A3 for the second and third one respectively. V stands for ‘verb’.

**Table 3:** All-new sentences: accented verbs.

	Arg 1	Arg 2	Arg 3	Verb	n
NDV	274	253		246	64
NAV	272	250		248	61
NDAV	291	263	249	246	45
NV	263			265	77

**Table 4:** All-new sentences: unaccented verbs.

	Arg 1	Arg 2	Arg 3	Verb	n
NDV	260	270		190	21
NAV	260	269		193	22
NDAV	277	244	271	187	45
NV	273			186	13

The following summarizing observations can be made from the graphs in Fig. 3 and the values of Tables 3 and 4.

First, the effect of downstep is visible in graphs a, c, d and in Table 3. A1, A2 and A3, when present, are clearly downstepped relatively to each other. An accented verb may continue the downstep (graph a) or be raised (graph c). But, except for the scaling of A1 and A2 in the NDAV condition, no downstep is found when the verb is unaccented and the preverbal argument is raised.

Second, some values are constant. The unaccented verbs have a value varying between 186 and 190Hz. We will see all along the paper that this value of the unaccented verb remains constant across all conditions and across all speakers. The pitch of the raised argument preceding the unaccented verb is also constant. We will see below that this is a target value and that is present in the narrow focus conditions as well.

Third, the starting point (height for the nominative) is higher in the longer sentence, a fact pointing to the existence of pre-planning (see section 5.4).

And the final remark: anticipating the discussion in sections 5 and 6, it can be observed that the patterns identified in the wide focused sentences arise as the results of different effects: downstep, H-raising and final  $F_0$  drop being the most prominent in the wide focus pattern.

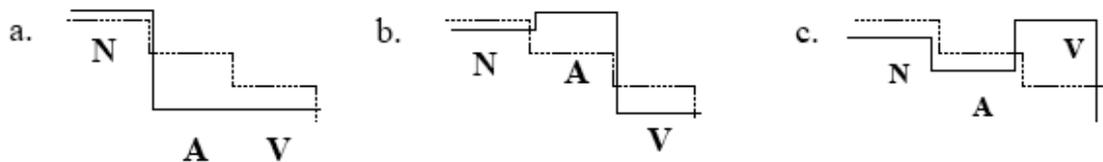
Before closing this subsection, the average values of all wide-focused realizations confounded are shown in table 5. Again, the effect of downstep is visible in Table 5, even though raising of the verb or of the preverbal argument are weighting in another direction.

**Table 5:** All-new sentences (all verbs)

	Arg 1	Arg 2	Arg 3	Verb	n
NV	264			254	90
NDV	271	257		232	85
NAV	269	255		233	83
NDAV	284	253	260	216	90

#### 4.1.2 Narrow focus

On the basis of the literature on the subject (see section 2), our expectations for the scaling of pitch accents in sentences containing a narrow focus are that first, a narrow focus raises a pitch accent in all positions, and second, givenness lowers prenuclear accents and suppress postnuclear ones. In Fig.4, the expectations related to the effect of narrow focus at different places in the sentence are displayed. The dotted line shows the regular downstep in an all-new sentence. This is the pattern that we assume to be the default (or unmarked) intonation. Fig.4a stands for an initial narrow focus on argument one (A1). A1 is higher than in the all-new configuration (because it is narrowly focused), argument two (A2) and the verb (V) are lower (because they are given and deaccented). Fig. 4b displays the expected effect of a narrow focus on a medial argument. A1 is expected to be lower than in the all-new configuration (because it is given and prenuclear), A2 is expected to be higher (it is narrowly focused), and V lower. Fig.4c shows the changes in pitch accents when the verb is narrowly focused. The pre-verbal arguments are lower than in the all-new condition and downstepped relatively to each other, since both are given, and there is an upstep on the verb.

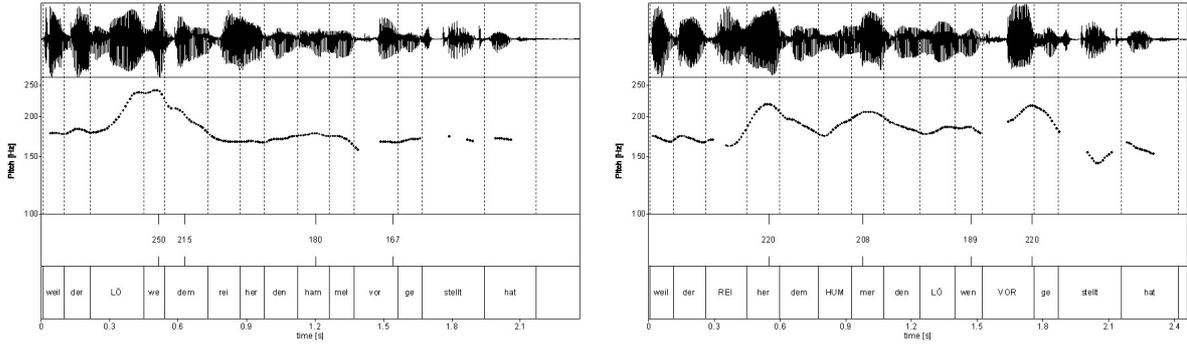


**Fig.4:** Effect of narrow focus (solid line) in comparison to a regular downstep pattern of the all-new sentence condition (dotted line). Fig. 4a on an initial argument, Fig. 4b on a non-initial argument, and Fig. 4c on a final verb.

Turning to the results of the experiments, the expectations related to the presence and scaling of pitch accents in the narrow focus configurations were confirmed. Much less variation in the accents was realized in sentences with a narrow focus than in those with a wide focus. A narrow focus systematically induced a bitonal accent on the narrowly focused argument or verb, as well as bitonal accents of all arguments preceding the narrow focus. All postnuclear constituents were systematically deaccented and consequently realized with a low and flat  $F_0$  contour.

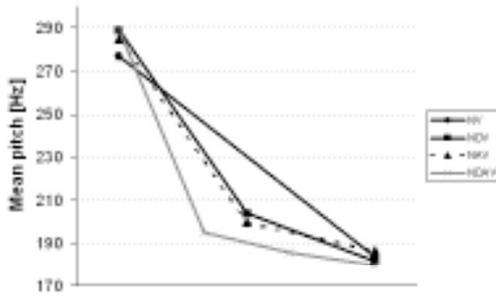
Narrowly focused sentences are illustrated in (7) and in Fig.5. In (7a), the first argument is focused, and all following arguments and the verb are unaccented. In (7b), the focus is on the verb. All preceding arguments have a pitch accent. Fig.5a shows a narrow focus on an initial nominative, and Fig.5b a narrow focus on the final verb. The postnuclear deaccented material in Fig.5a is much flatter than the unfocused prenuclear material in Fig.5b.

- (7) a. Weil (der HAMMEL)<sub>F</sub> dem Rammler den Hummer vorgestellt hat.  
 b. Weil der HAMMEL dem RAMMLER den HUMMER (VORGESTELLT)<sub>F</sub> hat.  
*'because the wether presented the lobster to the rabbit.'*



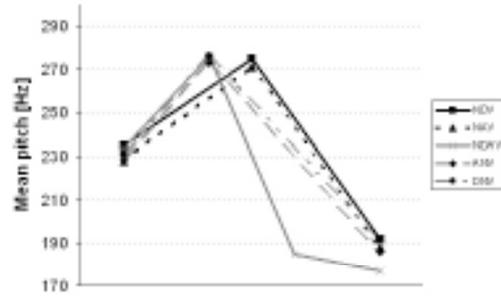
**Fig.5:** Narrow focus. Figure 5a (left panel) shows narrow focus on the initial constituent. Figure 5b (right panel) shows narrow focus on the verb.

In Fig.6, all graphs corresponding to the results of the sentences with a unique narrow focus (a A1, b A2, c A3, d V) are shown, accompanied by tables displaying the averaged high tones.



	NDV	NAV	NDAV	NV
Arg1	289	285	289	277
Arg2	204	200	195	
Arg3			186	
Verb	182	187	180	184

(a) narrow focus on the first argument



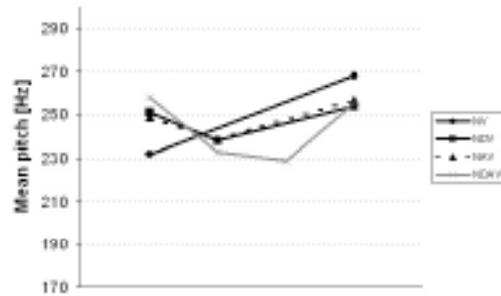
	NDV	NAV	NDAV	DNV	ANV
Arg1	235	228	234	231	229
Arg2	275	271	276	276	273
Arg3			185		
Verb	192	188	177	192	186

(b) narrow focus on the second argument



	NDAVN	NADV	ANDV	DANV	DNAV
Arg1	257	255	251	256	253
Arg2	229	230	224	227	226
Arg3	269	271	268	270	271
Verb	181	181	184	178	185

(c) narrow focus on the third argument



	NDV	NAV	NDAV	NV
Arg1	251	249	258	232
Arg2	238	239	233	
Arg3			229	
Verb	254	257	255	268

(d) narrow focus on the verb

**Fig. 6:** Representation of pitch peak realisation of narrow focus on different constituents.

Following observations are important for the remaining of this paper. A pitch accent on a narrow focus is always higher than the corresponding accent in a wide focus sentence. The following table sums up the values for A1, A2 and A3 in all-new conditions (AN) and narrow focus (NF) in comparable conditions.

**Table 6:** Comparison of the scaling of the arguments in all-new sentences (AN) with the corresponding in narrow focus (NF).

	$A1_{AN}$	$A1_{NF}$	$A2_{AN}$	$A2_{NF}$	$A3_{AN}$	$A3_{NF}$	$V_{AN}$	$V_{NF}$
NDV	271	289	257	275			232	254
NAV	269	285	255	271			233	257
NDAV	284	289	253	276	260	269	216	255
NV	264	277					254	268

A few summarizing remarks are necessary at this point. First, in line with all other arguments boosted by a narrow focus, a focused verb has a higher value when it is narrowly focused than when it is part of a wide focused sentence. It also has a much higher value when narrowly focused than when it is unaccented (see Table 4 for comparison). Compared to the first argument, the narrowly focused verb is only slightly higher or even lower than A1 (except for the NV pattern in which the verb is much higher than the unique argument), see Fig.6d.

Second, Table 7 illustrates that, across different conditions, the narrowly focused constituents are decreasing in height from the beginning to the end of the sentence. This is also an effect of downstep, as the constituents later in the sentences are preceded by accented constituents.

Third, when comparing the values of the first argument (A1) in different sentence conditions (first column of Table 7), it appears that the scaling of A1 varies as a function of the sentence condition. When comparing the values within a sentence condition (across rows), every argument is higher than the following one. The difference that was observed in the all-new sentences that are due to pre-planning (see below) have disappeared in the narrow focus condition.<sup>10</sup>

**Table 7:** Values of the narrowly focused constituents.

	$A1$	$A2$	$A3$	$V$
NDV	289	275		254
NAV	285	271		257
NDAV	289	276	269	255
NV	277			268

And finally, comparing further the wide focus results with those of this section, it can be noticed that A1 is higher when A2 is also given than when A2 is narrowly focused and A1 is immediately preceding the narrowly focused argument. This effect is analyzed as due to H-lowering, and is discussed in detail in section 5 (see table 13).

The remaining patterns are those in which more than one argument are focused, but the verb is not, see Table 8. In conditions NDV and NAV, two arguments are focused and in condition NDAV, three arguments. In the latter case, downstep applies between the first two arguments, followed by raising just before the verb. But this effect is not visible in the two-argument sentences, due to the effect of H-raising (see below).

<sup>10</sup> Thanks to Shin Ishihara for mentioning this fact.

**Table 8:** More than one argument is focused, the verb is given.

	A1	A2	A3	V	steps	n
<u>NDV</u>	264	270		188	+6 -82	78
<u>NAV</u>	257	269		189	+12 -80	82
<u>NDAV</u>	275	243	263	183	-32 +20 -83	89

#### 4.1.3 Statistical evaluation of the effect of focus structure

To evaluate the question of the impact of focus structure on the accent scaling a three-way ANOVA was performed with the  $F_0$ -maxima on each of the constituents being the dependent variable. The factors were (a) all-new sentences compared with a corresponding narrow focus, (b) the number of constituents from one (NV) to three (NDAV), and (c) position of the focus ranging from initial to third position (N2 stands for NV, N3 for NAV or NDV in any order, and N4 for NDAV in any order. P1, P2 and P3 stand for position 1 to position 3). The design of the analysis is shown in Figure 7.

Number of arguments			Positio n
n1P1	n2P1	n3P1	
	n2P2	n3P2	
		n3P3	

**Fig. 7:** The statistical design of the analysis of focus (n = no of arguments, P = position of the focused argument)

The first hypothesis to be tested is that the scaling of arguments in all-new sentences differs from the equivalent scaling in narrow focus structures. This hypothesis has been confirmed since we get a main effect of *focus structure* (factor (a)) ( $F(1,17) = 15.85$ ,  $p = 0.001$ ). We further observe a significant interaction of *focus structure* and *position* ( $F(1,17) = 11.50$ ,  $p = 0.003$  for the comparison of position one and two;  $F(1,17) = 12.93$ ,  $p = 0.002$  for the comparison of position two and three). Thus, the difference between all-new sentences and narrow focus depends on the position, the effect being stronger for position two than for position one and three. The interaction of *number of constituents* and *focus structure* showed no significant effect for one- and two-argument structures ( $p > 0.05$ ). Yet, a significant difference was found between two- and three-argument structures ( $F(1,17) = 9.51$ ,  $p = 0.007$ ) meaning that the scaling of longer sentences differs significantly between focus structures. Finally, the interaction of *focus structure*, and *number of constituents* and *position* taken together shows a significant effect ( $F(1,17) = 10.34$ ,  $p = 0.005$ ), which shows that the two factors have an impact on the scaling of the accents in different focus structures.

We may test the hypothesis that the scaling of narrow focused arguments in initial position differs as a function of the number of following arguments. The result is that we only find a significant effect for the comparison of one- and two-argument structures ( $F(1,17) = 21.02$ ,  $p = 0.000$ ) but not of two- and three-argument structures ( $p > 0.05$ ). Thus, the NV structures show a deviant pattern from NXV and NXYV structures concerning the factor *number of constituents*.

When considering the same comparison for narrowly focused arguments in second position, no significant difference ( $p > 0.05$ ) can be observed, which means that the height of the second argument is stable across conditions and speakers.

After having run paired-samples T-tests across speakers for the comparison of different word orders of the three-argument sentences, we could not find any significant effect on the scaling of the individual arguments (for all comparisons  $p > 0.05$ ).

## 4.2 Effect of givenness on pitch accents

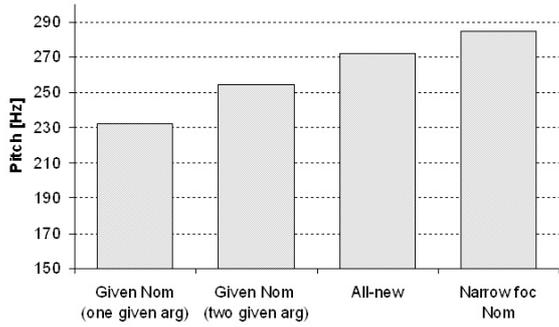
In the same way as focus was expected to raise the values of pitch accents, givenness is expected to lower them. As has been observed in the literature, prenuclear and postnuclear givenness have completely different effects. A prenuclear given constituent has a pitch accent, even if this pitch accent is comparatively low. A postnuclear given constituent is, by contrast, realized with a low and flat contour and is generally analyzed as being deaccented (see section 2 for discussion of the relevant literature). We inspect the pre- and postnuclear environment in turn. The relevant data can be gathered from Fig.6 or from the tables below.

To start with, compare the values of an initial nominative in four different contexts in Table 9 and Fig.8. The first column shows the value of a given nominative when it immediately precedes a narrow focus. In the second column, the given nominative precedes another given constituent. The third column shows the values of a nominative in an all-new sentence, and the last one when it is a narrow focus itself. As can be seen, these values increase significantly from column to column, showing a four-way contrast in the height of this initial constituent. Paired samples T-tests comparing the scaling of the initial nominative in different focus structures reveal that given constituents differ significantly from the corresponding nominatives in all-new sentences ( $df = 17$ ,  $t = 9.53$ ,  $p = 0.000$  for nominatives followed by an accented constituent;  $df = 17$ ,  $t = 4.51$ ,  $p = 0.000$  for nominatives followed by a further given constituent). The same holds for the comparison with a narrow focused nominative ( $df = 17$ ,  $t = 12.36$ ,  $p = 0.000$ ;  $df = 17$ ,  $t = 7.02$ ,  $p = 0.000$ ). Comparing between the scaling of the two initially given nominative structures, we observe a significant effect as well ( $df = 17$ ,  $t = -8.50$ ,  $p = 0.000$ ).

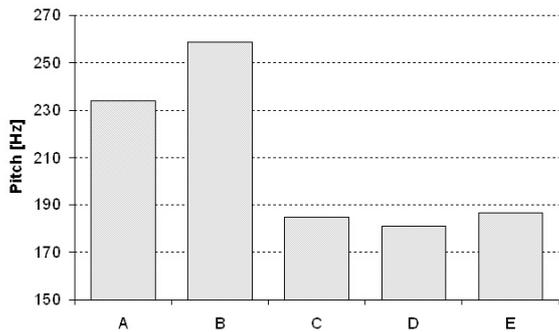
Clearly, a focus-given distinction cannot account for this pattern. Even if all-new, narrow focus and given are distinguished from each other, there still remain an unaccounted fourth value. The unexpected fourth value comes from the following observation: a given nominative preceding a narrowly focused constituent (first column) is lower than a nominative followed by another given constituent (second column). This effect is attributed below to H-lowering (see section 5.4).<sup>11</sup> Preplanning has nothing to say here, because we are not comparing sentences with different length, but only sentences in which the first argument has a different role in term of focus-given.

<sup>11</sup> More values illustrating that prenuclear accents are preserved in a given sentence, although they are much lower than in the wide focus sentences are summed up in the next table for A2 and A3.

	A2 <sub>WF</sub>	A2 <sub>Giv</sub>	A3 <sub>WF</sub>	A3 <sub>Giv</sub>
(NAV/NDV)	256	236		
(NDAV)	253	227	260	229
(NV)				



**Fig. 8:** The initial nominative in different environments.



**Fig. 9:** Effect of givenness on the verb. A – All new condition, B – Narrow focus on the verb, C – Narrow focus on pre-verbal constituent, D – Narrow focus on pre-verbal constituent (pre-verbal constituent and verb are unaccented (given)), E – All pre-verbal constituents are focused (accented).

**Table 9:** Values of initial nominative in different information structures.

Given Nom: one given argument		Given Nom: two given arguments		All-new		Narrow focus on Nom	
<u>N</u> V	232	—		<u>N</u> V	264	<u>N</u> V	277
<u>N</u> DV	235	<u>N</u> DV	251	<u>N</u> DV	271	<u>N</u> DV	289
<u>N</u> AV	228	<u>N</u> AV	249	<u>N</u> AV	269	<u>N</u> AV	285
<u>N</u> DAV	234	<u>N</u> DAV	257	<u>N</u> DAV	284	<u>N</u> DAV	289
		<u>N</u> ADV	255				
		<u>N</u> DAV	258				

In the postnuclear position, accents are deleted. Fig.9 compares the value of the verb in different environments, also given in Table 11. The first column shows the height of the verb in an all-new sentence. The second one stands for its values in a narrow focus environment. As can be seen by comparing columns 1 and 2, the verb has a higher pitch when narrowly focused than when included in an all-new pattern. The third column shows the height of the verb when a narrow focus immediately precedes the verb. The fourth column shows the scaling of the verb when the non-immediately preceding constituent is given as well. And finally, the fifth column shows the scaling of the verb when all constituents but the verb are focused.

When narrowly focused, the verb is scaled around 258Hz. In the postnuclear position, the verb has a very stable value, as can be gathered from columns 3 to 5 from table 11.

**Table 11:** Values of unaccented and accented verbs

All-new	Narrow focus on the verb	Narrow focus on pre-verbal constituent	Narrow focus on pre-preverbal constituent	All pre-verbal constituents are focused
<u>NV</u> 254	<u>NV</u> 268	<u>NV</u> 184	—	—
<u>NDV</u> 232	<u>NDV</u> 254	<u>NDV</u> 192	<u>NDV</u> 182	<u>NDV</u> 188
<u>NAV</u> 233	<u>NAV</u> 257	<u>NAV</u> 188	<u>NAV</u> 187	<u>NAV</u> 189
<u>NDAV</u> 216	<u>NDAV</u> 255	<u>NDAV</u> 181	<u>NDAV</u> 177	<u>NDAV</u> 183
		<u>ANV</u> 186	<u>NDAV</u> 180	
		<u>DNV</u> 192		
		<u>NADV</u> 181		
		<u>ANDV</u> 184		
		<u>ADNV</u> 182		
		<u>DNAV</u> 185		
		<u>DAN</u> 178		

Testing the hypothesis that unaccented verbs differ in their scaling from accented verbs we run paired samples t-tests for the given verbs with the focused ones individually. A given verb when preceding an accent is scaled significantly lower than in an all-new sentence ( $df = 17$ ,  $t = 10.23$ ,  $p = 0.000$ ) and in narrow focus ( $df = 17$ ,  $t = 18.19$ ,  $p = 0.000$ ). Given verbs do not differ significantly from each other as a function of different preceding accent position (paired samples t-tests for all comparisons  $p > 0.05$ ).

Postnuclearly, medial values are lower when the constituents are given than when they are new. This is shown in Table 12, for some of the conditions. The information status of the initial constituent is kept constant, and the status of the medial argument is changed. It is given in the first column and focused in the third one. The table shows that the value of the accented argument in the right part of the table is higher before an unaccented verb (first two rows) than before a following accented argument or verb (last two rows). The differences between the values of a focused pitch accent and a given one varies between 48 and 78Hz.

**Table 12:** Values of medial arguments.

	unaccented		accented	differences
<u>NAV</u>	200 (A)	<u>NAV</u>	269 (A)	70
<u>NDV</u>	204 (D)	<u>NDV</u>	270 (D)	62
<u>NDAV</u>	195 (D)	<u>NDAV</u>	243(D)	48
<u>NDAV</u>	185 (A)	<u>NDAV</u>	263 (A)	78

Following observations arise from the discussion in this subsection. First, the prenuclear given material is lower than corresponding accents in the wide focus condition. A prenuclear given argument is even lower when immediately preceding a narrow focus than when it is separated from the narrow focus by another given argument (due to H-lowering effect, see section 5.4).

Second, as can be seen from the values summed up in Fig.6, when the third argument is a narrow focus, the first two arguments are in a downstep relationship to each other, and the steps have approximately the same size (around 27 Hz). The value of the third argument, the focused one, is higher than the second one by 42Hz. The difference varies between 40 and 44Hz, and is similar to the difference between prenuclear and nuclear argument when A2 is narrowly focused. The value of the accented argument is higher than those of the initial argument by approximately 15Hz.

Third, focused material is much higher than postnuclear material.

Finally, when a second (and third) postnuclear constituent is present, downward steps are still realized, even though they become much smaller as the end of the sentence is approaching. See the values in Fig.6a, and marginally 6b. We do not try to interpret this decrease in pitch in the postnuclear stretch of the sentence in this paper, since this could be the topic of another study, but notice that this is an issue to be investigated.

### 4.3 Effect of focus on duration

The results for duration can be summed up easily. Narrowly focused constituents are significantly ( $F(1,17) = 20.775, p < 0.000$ ) longer than non-focused ones, a result which is shown in Fig.10. A second result is also apparent from this figure: nominative arguments are shorter than their accusative and dative correspondents. This difference may reflect a difference in their position in the sentence since a nominative is more often initial than the other constituents. But we do not seek to interpret this effect. The longer duration of the verb correlates with the fact that it is quadrisyllabic (except *vor.ge.stellt* 'introduced'), as opposed to the arguments which are always trisyllabic (article + noun), and is thus irrelevant.

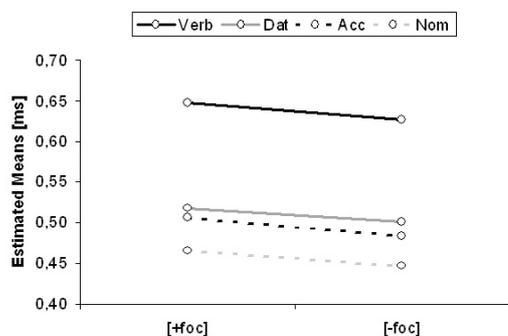


Fig. 10: Effect of focus on duration.

A further interesting (and unexpected) effect of duration correlates with the number of constituents that a sentence contains. Fig.11 shows that the duration of an initial nominative increases significantly ( $F(2,34) = 21.60, p = 0.000$ ) when followed by one, two and three constituents (verb included). It is conspicuous that the realization of the nominative takes more time when more constituents are to follow. This surprising effect could correlate with preplanning, though in a poorly understood way. A longer stretch of arguments causes the speaker to slow down her tempo.<sup>12</sup>

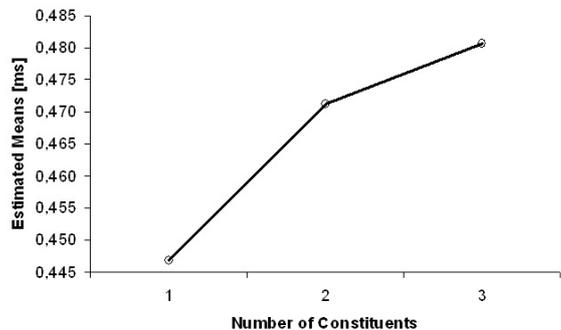
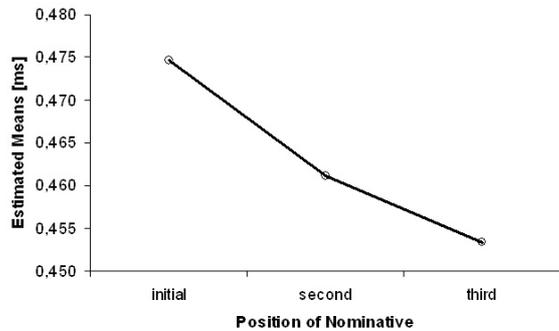


Fig. 11: Effect of number of constituents on the duration of the nominative.

<sup>12</sup> Thanks to Sabine Zerbian for attracting our attention to the potential correlation between longer duration and preplanning.

Finally, duration does not vary significantly ( $F(2,34) = 8.60, p > 0.05$ ) as a function of the word order. This holds for the nominative in the initial, second and third position in the sentence. We hypothesize that the small decrease in length observed in Fig.12 cannot be an artifact of the changed word order, because a late nominative corresponds to a marked word order, so that longer duration would be expected as a result. But further study is necessary to confirm this hypothesis.



**Fig. 12:** Effect of position of the nominative on duration.

#### 4.4 Word order

Word order did not affect any of the values calculated in this study, as was shown for duration just above. As far as pitch is concerned, we first compare the accents of the arguments in the unmarked patterns  $\underline{N}\underline{A}\underline{V}$  and  $\underline{N}\underline{D}\underline{V}$  with those in the marked patterns  $\underline{A}\underline{N}\underline{V}$  and  $\underline{D}\underline{N}\underline{V}$ . Sentences with the same arguments but in a different order, do not differ in the height of their arguments. In a sentence with a nominative and an accusative, the first non-focused and preverbal focused argument has the same height respectively, regardless of whether it is the nominative or the accusative which is found there. And the same observation holds for a sentence with a nominative and a dative. These results can be gathered from the following table.

In the longer conditions, again no difference can be found in the height of the accents. The values are similar in unmarked and in marked word orders.

In sum, no significant effect of word order was found, neither in the focused constituent nor in the given ones. But, as this can be an artifact of the experiment in which informants spoke in a rather mechanical way, we consider this issue as unresolved.

**Table 13:** Effect of word order on a sentence with two arguments.

	A1	A2	V
$\underline{N}\underline{A}\underline{V}$ (unmarked)	228 (N)	271 (A)	188
$\underline{N}\underline{D}\underline{V}$ (unmarked)	235 (N)	275 (D)	192
$\underline{A}\underline{N}\underline{V}$ (marked)	229 (A)	273 (N)	186
$\underline{D}\underline{N}\underline{V}$ (marked)	231 (D)	276 (N)	192

**Table 14:** Effect of word order on a sentence with three arguments

	A1	A2	A3	V
$\underline{N}\underline{D}\underline{A}\underline{V}$	257	229	269	181
$\underline{N}\underline{A}\underline{D}\underline{V}$	255	230	271	181
$\underline{A}\underline{D}\underline{N}\underline{V}$	255	227	271	182
$\underline{A}\underline{N}\underline{D}\underline{V}$	251	224	268	184
$\underline{D}\underline{A}\underline{N}\underline{V}$	256	227	270	178
$\underline{D}\underline{N}\underline{A}\underline{V}$	253	226	271	185

## 4.5 Summary

Summing up this section, effects of information structure on the height of tones could be clearly identified. Likewise, duration is affected by information structure (focused constituents are longer than unfocused ones), but no effect of word order was found.

In an all-new sentence, two different realizations could be identified. Either an unmarked downstep contour, or an upstep on the last accented constituent, the preverbal argument or the verb. When a constituent is narrowly focused, several effects are exerting an influence on the tonal pattern of the entire sentence. First the pitch value of the narrow focused constituent is raised. Second, the preceding arguments are realized with lower pitch accents than in conditions in which these arguments are focused. Third, postnuclear accents are suppressed.

But not all changes in height can be explained by the effect of information structure. Specifically, the relative scaling of pitch accents depends on other effects which are analyzed in the next section as tonal effects.

## 5. Tonal effects

Additionally to the effects of information structure, purely tonal effects are also playing a substantial role in scaling tones relatively to each other. First, downward trends are examined: downstep in 5.1 and final  $f_0$  drop in 5.2. The phenomenon of H-raising is the subject of section 5.3, and H-lowering of section 5.4. We briefly give an account of preplanning in section 5.5, as it was observed in our data, and sum up the tonal effects in 5.6.

### 5.1 Downstep

Regular downstep takes place in approximately half of the all-new sentences ( $N = 159$ , 45.7%), as discussed in section 4.1.1. We assume this pattern as the default tonal realization in which unmarked prosody causes relatively small but regular descending steps, between 17 and 29 Hz.

Table 5, repeated as Table 15, shows downstep in wide focus sentences, all patterns confounded.

**Table 15: Downstep in all-new sentences (all verbs)**

	A1	A2	A3	V	n
NDV	271	257		232	85
NAV	269	255		233	83
NDAV	284	253	260	216	90
NV	264			254	90

A further environment for downstep is a sequence of two given arguments followed by a focused one. In this case, the given arguments also show downstep, as can be seen from Table 16.

**Table 16:** Downstep in initial given arguments preceding a narrow focus

	1st	2nd	3rd	steps	n
<u>NAV</u>	249	239		-10	89
<u>NDV</u>	251	238		-13	90
<u>NDAV</u>	257	229		-28	90
<u>NADV</u>	255	230		-25	88
<u>ADNV</u>	255	227		-28	90
<u>ANDV</u>	251	224		-27	90
<u>DANV</u>	256	227		-29	90
<u>DNAV</u>	253	226		-27	90
<u>NDAV</u>	258	233	229	-25,-4	84

In Table 16, the downward steps are about 12 Hz when only two arguments precede the accented verb (two-argument sentences). When the third argument is focused, the downward steps are larger: around 27 Hz. This difference can be explained by the fact that in the first two configurations, it is the verb which carries the narrow focus, and as we saw above, the values of the pitch accent on a narrowly focused verb are lower than the ones on a preceding constituent.

The last configuration in which we find downstep is illustrated in Table 17, when only the arguments are focused to the exception of the verb. In the condition NDAV, the difference in height between A1 and A2 is of 31Hz. A3 is upstep.

**Table 17:** Downstep in initial new arguments preceding an unaccented verb.

	1st	2nd	3rd	V	steps	n
<u>NDAV</u>	275	243	263	183	-32 +20 -83	89

In short, downstep features regularly in a sequence of two or more arguments when they have the same informational status, that is when they are equally focused or when they are equally given.

## 5.2 Final $f_0$ drop

A much larger downward trend is observable between the pre-verbal accented argument and the final unaccented verb than downstep would lead us to expect. At the place where the last accent is realized, two effects co-occur: the H tone is raised (cf. next section) and a considerable subsequent drop in  $f_0$  is taking place. The final  $f_0$  drop between the preverbal accented argument and the unaccented verb betrays in average 80 Hz.

Since the value of the final unaccented verb has a stable value at around 190Hz, and since the final drop is also constant with 80-85Hz, it follows that the value of the preverbal accented argument is also constant. And indeed, this is exactly what we find.

A final drop is also present between a narrow focus and the immediately following unaccented constituent. Interestingly, it is not the low value reached by the final drop which is constant but the drop itself. As before, it has a constant value of around 84Hz, but it reaches different values depending on whether it is the first, second or third argument which is narrowly focused. At the end of the sentence, the value of 190Hz or less is reached in all cases, but sometimes it needs more than one decreasing steps to obtain (see the values in Fig.6a-c).

Paired samples T-tests on the comparison of the final drop across different conditions show no significant difference between narrow focus in different positions (for all comparisons  $p > 0.05$ ). However, comparing the final drop between the all-new

condition with that of the narrow focus conditions, we observe a significant difference in that the final drop is significantly larger for narrow focus ( $df = 17, t = 6.58, p = 0.000$  for all-new vs. initial narrow focus;  $df = 17, t = 7.79, p = 0.000$  for all-new vs. second argument focus;  $df = 17, t = 7.60, p = 0.000$  for all-new vs. third argument focus). The difference is due to the fact that the all-new sentences had different realizations, only some of them involving final  $F_0$  drop.<sup>13</sup>

In the following graph, [+accent] indicates the value of the last accent in the focused material, and [-accent] is the value of the first unaccented constituent.

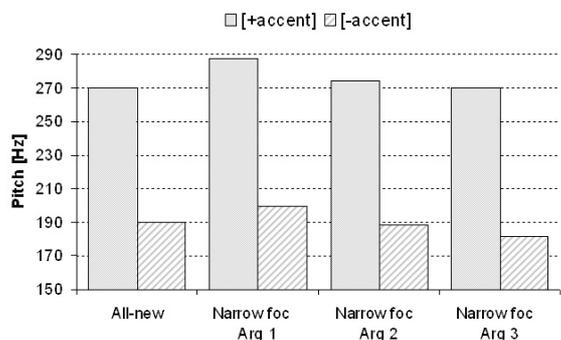


Fig. 13: Final drop

Table 18: Final drop in different conditions.

Final drop after		Accented arg.	Unacc. const.	Final Drop
last	<u>NAV</u>	269 (A)	193 (V)	76
argument	<u>NDV</u>	270 (D)	190 (V)	80
(all-new)	<u>NDAV</u>	271 (A)	187 (V)	84
initial	<u>NAV</u>	285 (N)	200 (A)	85
focus	<u>NDV</u>	289 (N)	204 (D)	85
	<u>NDAV</u>	289 (N)	195 (D)	94
second	<u>NAV</u>	271 (A)	188 (V)	83
argument	<u>ANV</u>	273 (N)	186 (V)	87
focus	<u>NDV</u>	275 (D)	192 (V)	83
	<u>DNV</u>	276 (N)	192 (V)	84
	<u>NDAV</u>	276 (D)	185 (A)	91
third	<u>NDAV</u>	269 (A)	181 (V)	88
argument	<u>NADV</u>	271 (D)	181 (V)	90
focus	<u>ANDV</u>	268 (D)	184 (V)	84
	<u>DANV</u>	270 (N)	178 (V)	92
	<u>DNAV</u>	271 (A)	185 (V)	85

<sup>13</sup> For the analysis of the all-new sentences, we only calculated the final drop for sentences containing an unaccented verb – these are 101 instances of a total of 348 all-new sentences. Recall from Fig. 2 that 41% of these had a regular downstep pattern before the drop while 59% showed upstep of the preceding argument (as is similar to a narrow focus).

### 5.3 H-Raising

According to Laniran & Clements (2003), Xu (1999) among others, H-raising characterizes a raised value of a H preceding a low tone in tone languages. In German, H-raising is observed, as well, but only on the last accented word, just before the final drop. It is not limited to the preverbal accent, but it raises the values of an accent when it is directly before an unaccented one, disregarding its position in the sentence. Table 19 compares the values of a high tone when it precedes a further accented argument with its value when it precedes an unaccented argument.

From all relevant comparisons in Table 19, it appears that an argument is lower when it precedes an accented argument (second column) than when it precedes an unaccented one (fourth column). In average, the difference is of 17 Hz. There are only two places where the difference is not significant. The first one is N of NDAV, thus on a nominative followed by two more arguments (only 5Hz difference). We will see below that this is a place where the highest value is reached. Because of preplanning, speakers start higher when more arguments are to follow. In other words, the voice reaches a ceiling in this case, independently of the status of the following argument. The second place where the difference is not significant is A3 in NDAV.

**Table 19:** Effect of H-raising on an accented argument.

	before [+foc]		before [-foc]
<b>A1</b>			
<u>NV</u>	264	<u>NV</u>	277
<u>NDV</u>	271	<u>NDV</u>	289
<u>NAV</u>	269	<u>NAV</u>	285
<u>NDAV</u>	284	<u>NDAV</u>	289
<b>A2</b>			
<u>NDV</u>	257	<u>NDV</u>	270
	257	<u>NDV</u>	275
<u>NAV</u>	255	<u>NAV</u>	269
	255	<u>NAV</u>	271
<u>NDAV</u>	253	<u>NDAV</u>	276
<b>A3</b>			
<u>NDAV</u>	260	<u>NDAV</u>	263
	260	<u>NDAV</u>	269

The results of table 19 for A1, A2 and A3 are summed up in Fig. 14 where all arguments in an all-new sentence (uniformly grey block) are compared their narrowly focused counterparts when no further accent is following (shaded block). The scaling of a narrow focus is much higher as compared to the equivalent argument in the all-new condition. Paired samples T-tests prove this difference as significant ( $df = 17$ ,  $t = -4.12$ ,  $p = 0.01$  for the first argument;  $df = 17$ ,  $t = 5.03$ ,  $p = 0.000$  for the second argument).

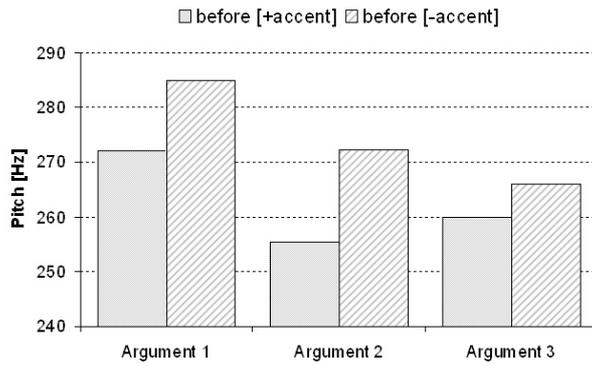


Fig. 14: H-raising.

### 5.4 H-Lowering

L-lowering, which can be understood as the reverse effect of H-raising, is also observable in our data. A high tone is lowered when it immediately precedes a raised tone. Several data sets show this effect.

First, as shown in Table 9, the lowest values for an initial nominative are those in which it immediately precedes a narrow focus (leftmost column). When the second argument is given as well (second column), the values of the nominative are higher by ca. 20 Hz. This difference cannot be explained by any other effect than a purely tonal one. In particular, it cannot be explained if only information structural effects are taken into account, since the given status is identical in both cases.

Second, table 20 compares the height of arguments directly preceding a narrowly focused argument with similar ones in a wide focused pattern.

**Table 20:** Lowering of a high tone (bold face) preceding a narrow focus.

				differences
<u>NAV</u>	A= 239	<u>NAV</u>	A= 255	11
<u>NDV</u>	D= 238	<u>NDV</u>	D= 257	15
<u>NDAV</u>	D= 229	<u>NDAV</u>	D= 253	34
<u>NDAV</u>	A= 229	<u>NDAV</u>	A= 260	16

The values of the argument immediately preceding a narrow focus (second column in table 20) are again surprisingly stable, and systematically lower than the ones they have in a wide focused sentence (fourth column).

Even if it is assumed that givenness is partly responsible for the low values in the second column of table 20, such an interpretation is not available for the third set of data: an argument preceding a H-raised tone in a wide focused sentence. In this case, as well, an argument preceding a raised value is lowered as compared to a downstep pattern. Since the information status of all constituents is identical in each case, it must be assumed that it is the tonal make-up of the phrase which is responsible for the difference. The values are given in tables 21 and 22.

In the last two cases, the differences is again of 17 Hz or more.

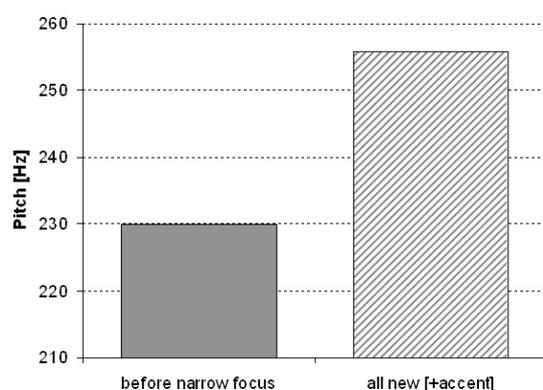
Paired samples T-tests show that a constituent prior to an accented constituent is significantly lower scaled than a corresponding constituent in an all-new sentence ( $df = 17$ ,  $t = 7.06$ ,  $p = 0.000$ ).

**Table 21:** Lowering of a high tone preceding an upstepped verb.

	Verb is upstepped	Downstep pattern	acc. V, reg. downstep	all-new	differences
<u>NAV</u>	A= 245	<u>NAV</u>	A= 251	255	17
<u>NDV</u>	D= 242	<u>NDV</u>	D= 259	257	17
<u>NDAV</u>	A= 230	<u>NDAV</u>	A= 258	260	28

**Table 22:** Lowering of a high tone preceding an upstepped argument.

Preverbal arg. is upstepped		Downstep pattern	acc. V, reg. downstep	all-new	differences
<u>NAV</u>	N= 260	<u>NAV</u>	N= 276	269	16
<u>NDV</u>	N= 264	<u>NDV</u>	N= 283	271	19
<u>NDAV</u>	D= 257	<u>NDAV</u>	D= 275	253	18

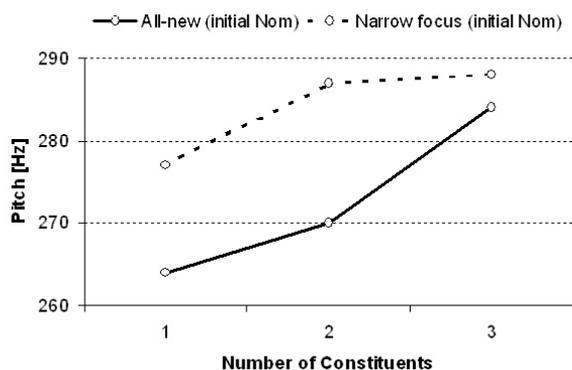
**Fig. 15:** H-lowering.

## 5.5 Preplanning

The last question addressed in this section concerns preplanning. Is it the case that an initial Nominative in an all-new sentence is higher when more arguments are to come? This question must be answered affirmatively. The average values of an initial nominative in a whole-focused sentence are clearly higher when more arguments are to come. Table 23 shows the values with all-new pattern. In case of a narrow focus on the nominative, the difference is not significant. Here it is to be assumed that speakers have a top line which cannot be raised. The boosting effect of a narrow focus raises the value of an accent so that this top value is reached in all situations. But still the difference between the shorter and the longer sentences shows an increase of 10 Hz. The statistical significance of the comparison in Table 23 has already been made in section 4.1.3 above.

**Table 23:** Values of initial accented nominative as a function of number of constituents.

	All-new (N)	Narrow focus (N)
NDV	271	289
NAV	269	285
NDAV	284	289
NV	264	277



**Fig. 16:** Preplanning

## 5.6 Summary

This section has introduced effects which are to be interpreted as the result of the interactions of tones with each other. We have seen that downstep, final  $f_0$  drop, H-raising and H-lowering are effects regulated by the tonal pattern itself. The question that remains to be discussed is how purely tonal effects like these are to be integrated in the phonological representation of German, an intonation language. The question has an important theoretical relevance, as it impinges on the representation of tones in general.

## 6. Discussion

In this last section, we discuss the results of the experiment in the framework of a larger phonological context. It is first proposed that prosodic phrasing is not changed by differences in information structure (6.1), but that only pitch scaling is involved in the phonological expression of focus and givenness (6.2 and 6.3). It is shown that the exact melody of sentences is the product of an interaction between information structure and tonal effects. In the course of the discussion, alternative explanations are addressed and rejected, as they cannot account for the phenomena identified here as smoothly as our proposal.

### 6.1 Prosodic phrasing

Sentences consisting of only one intonation phrase, like those used in the experiment, are divided into a number of prosodic phrases according to the constituents they contain. Prosodic phrasing is best analyzed as a structure defined exclusively by syntax, i.e. as a mapping from a syntactically organized strings of words to a phonological representation. Whether the head of a prosodic phrase is accented or not is a different matter, which depends only in part on the syntactic structure itself, and in part on the information structure. Length, weight, rate of speech, style or lexical content may also influence the placement of pitch accents, but they are not considered here.

The sentences in (15) to (17) stand for all-new sentences. All our sentences are verb-final embedded clauses, in which all arguments are in the middle field. The upshot of this syntax is that the subject is behaving like the other arguments, and is not separate syntactically or prosodically from the remaining of the clause. The predicate is phrased together with the immediately preceding argument, and all other arguments are forming

their own phrase (see Fuchs 1976, Jacobs 1993 and Gussenhoven 1992 for proposals taking this into account). In the following examples, a subscripted P indicates prosodic phrasing and small caps highlight the expected accents. (15a and b) differ only in the accent on the verb (see above).

- (15)a. [Weil der HAMMEL]<sub>P</sub> [den HUMMER eingeladen hat]<sub>P</sub>  
 b. [Weil der HAMMEL]<sub>P</sub> [den RAMMLER EINGELADEN hat]<sub>P</sub>  
 c. [Weil der RAMMLER]<sub>P</sub> [dem REIHER ]<sub>P</sub> [den HUMMER vorgestellt hat]<sub>P</sub>

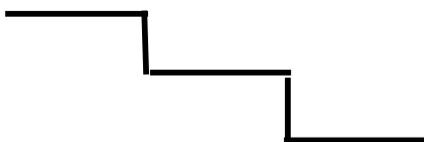
We propose that information structure does not change the prosodic structure, which is driven by syntax only.<sup>14</sup> In (16) different information structures are shown, but the prosodic phrasing remains the same. Subscripted F stands for focus, and G for givenness. In (16c), the prenuclear constituent *Rammler* is accented, but the postnuclear one, *Hummer*, is not.

- (16)a. ([Weil der HAMMEL]<sub>P</sub> [den RAMMLER eingeladen hat]<sub>P</sub>)<sub>F</sub>  
 b. ([Weil der HAMMEL]<sub>P</sub>)<sub>F</sub> ([den Rammler eingeladen hat]<sub>P</sub>)<sub>G</sub>  
 c. ([Weil der RAMMLER]<sub>P</sub>)<sub>G</sub> ([dem REIHER ]<sub>P</sub>)<sub>F</sub> ([den Hummer vorgestellt hat]<sub>P</sub>)<sub>G</sub>

## 6.2 Register and reference lines

If, in our material, prosodic phrasing is insensitive to information structure, the question arises as to what kind of influence the information structure exert on prosody. The answer has been given all along this paper: tonal register, and pitch accent scaling are the reflexes of information structure. In this section and the next one, only the high part of pitch accents are considered. The low part has been shown in the literature to be less sensitive to downstep, and scaling issues in general (see Ladd 1984, Grabe 1998, Truckenbrodt 2002 among others).

In the unmarked case, each prosodic phrase (henceforth p-phrase) in an intonation phrase (i-phrase) has its own register, and the registers are organized in a downstep fashion (see Clements 1990, Ladd 1992, van den Berg et al. 1992, Truckenbrodt 2004 and Féry & Truckenbrodt 2005 for this understanding of downstep). The register lines can be understood as intervals of the speaker voice range, in the limits of which pitch excursions are scaled. In a sequence of p-phrases in an i-phrase, every p-phrase is by default lowered relatively to the preceding p-phrase.



**Fig. 17:** Downstep of reference lines in a sequence of p-phrases

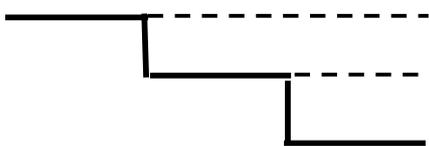
Our view of downstep follows the tradition introduced by Pierrehumbert (1980) and Liberman & Pierrehumbert (1984) for intonation languages, in which a medial L tone affects a following H in such a way that it is lower than a preceding H (see also Truckenbrodt 2004 for German). This effect takes the form of a left-to-right anticipatory

<sup>14</sup> There is a sense in which prosodic structure is changed by information structure through the intermediary of syntax: topicalization, object preposing, right and left dislocation and the like, create new prosodic phrases, and syntax structure is changed accordingly. But this is a different matter, and this paper only examines prosodic phrasing inside a single intonation phrase, when word order is changed minimally.

assimilation. It is assimilatory because a low tone lowers a following high tone, thus rendering it more similar to itself.

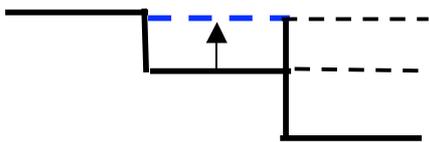
Our analysis differs from the ones just mentioned in a single respect. We assume that downstep takes place across p-phrases as well as inside of p-phrases. Since our data contain p-phrases with maximally one tone, only downstep across p-phrases are of interest here. We analyze the high part of a rising accent as the boundary tone of the p-phrase it ends. The high parts of the rising contour ( $H_p$ ) are maximally as high as the reference line of their own domain. Every non-final accent is represented phonologically as  $L^*H_p$  and the final accent is a falling tone, that we represent with a high starred tone followed by the low boundary of the i-phrase,  $H^*L_i$ . Downstep creates new (and downstepped) reference lines. The first  $L^*$  tone of a p-phrase lowers the following  $H_p$ , and it is this which creates the staircase pattern of Fig.17.

The register lines created at each p-phrase, are available until the end of the i-phrase, see Fig.18.



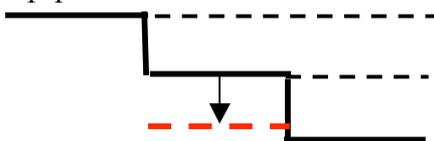
**Fig. 18:** Reference lines are available until the end of the i-phrase

A narrow focus raises the reference line of its corresponding domain, as shown in Fig.19.<sup>15</sup> In other words, a narrow focus expands the available range of accent and tone. In this paper, only cases in which the corresponding domain of a focus is equivalent to a prosodic phrase are considered, but this one-to-one-correspondence is not necessary: a focus domain may be larger or smaller than a prosodic phrase (see Ishihara 2003, 2006 for Japanese). Crucially, a raised reference line affects all focused material (but not an all-new sentence).



**Fig.19:** Boosting of the reference line due to narrow focus

Similarly, givenness lowers the register of the prosodic phrase corresponding to the given constituent, as illustrated in the following figure. Thus givenness reduces the  $f_0$  range available. Again, lowering can affect more than one p-phrase, or even just part of a p-phrase.

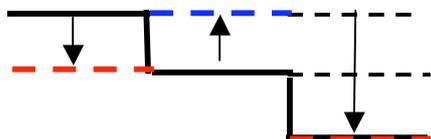


**Fig.20:** Lowering of the reference line due to givenness

<sup>15</sup> We conjecture that the value of the raised (and lowered) reference lines target values that are variable according to the amount of newness or givenness of the material. A contrastive focus raises the reference lines more than a normal informational focus. As for givenness, a Second Occurrence Focus lowers the reference line less than a backgrounded element (see Féry & Ishihara 2005 for Second Occurrence Focus, and Baumann 2006 for different kinds of givenness).

After the last accent of the sentence, the reference line of the remaining i-phrase is lowered to a minimum. The absence of postnuclear accent reflects this extreme compression of the reference line. As the register available is reduced to a minimum, there is no space left for pitch accents.

The effect of givenness, narrow focus and postnuclear compression may combine in a single intonation phrase, as shown in the following figure.



**Fig.21:** Postnuclear total compression of the reference lines

The reference lines are thus the phonological reflexes of information structure. An all-new sentence has downstepped reference lines, givenness lowers them and a narrow focus has the effect of raising the reference line. Postnuclearly, the reference lines are set to zero. Their role is to define register domains for pitch accents. In a more abstract phonological model, reference lines can be conceived of as higher register lines in a tone tree, as has been proposed proposed by Clements (1990) and Ladd (1986), as well as by Féry & Truckenbrodt (2005).

As shown in section 5, there is more variation in the scaling of individual high tones than predicted by this simple raising and lowering of reference lines. The second kind of effects are purely tonal. They come from the interaction of the tones with each other within the limits imposed by the reference lines.

A tone may act on an adjacent one in an assimilatory or dissimilatory way (Xu 1997). We already saw that downstep is an assimilatory effect that a low tone exerts on a following high tone, creating in this way a stepwise decrease in the value of the high tones. By contrast, both H-raising and H-lowering are regressive effects, and both are dissimilatory. An L tone raises the value of a preceding high tone, whereas an H tone lowers a preceding H tone. In both cases, a tone is becoming less similar to the following tone.

## 6.2 Final $f_0$ drop and H-raising

Let us first examine Final  $f_0$  drop and H-raising, which involve the i-phrase-final sequence  $H^*L_i$ . A final low tone ( $L_i$ ) reaches very low regions and at the same time, it raises the value of an immediately preceding  $H^*$ . The process takes place in two steps. First, the final low tone,  $L_i$ , is scaled at the bottom line of the voice register, or near to it, if it is not the last accentable constituent of the sentence. This is due to the drastic lowering of the reference line in the postnuclear position. This  $L_i$  is a boundary tone, associated with the end of the i-phrase in a declarative sentence. It is also aligned with the syllable following the nuclear accent (see Gussenhoven 2004 for an OT analysis along these lines for English), creating in this way a postnuclear flat and low contour.<sup>16</sup> The second effect is the increase in pitch that  $L_i$  is exerting on the preceding  $H^*$ . It raises its value independently of what tones precede the  $H^*$ . We saw that the value of

<sup>16</sup> This may be a simplification, due to the observation that a non final postnuclear word may be followed by decreasing steps, as shown in section 4.2, but it does not bear on the argument of this section.

the raised H\* is dependent on Final  $f_0$  drop, as it is the difference between H\* and  $L_i$  which seems to matter.

This analysis of H-raising leans on studies uncovering a similar effect in a number of tone languages. Xu (1997) examines anticipatory and carry-over effects of Mandarin tones, and finds that a high tone is higher when the following tone is low. He concludes that a dissimilatory effect is at play. Similarly, Laniran & Clements (2003) finds a dissimilatory H-raising effect in Yoruba: A low tonal onset raises the  $f_0$  of the preceding high tone. Gandour et al. (1994) also find an H-raising effect in Thai. A strict implementation of the tone sequence analysis of intonation predicts such tonal effects to be universal, in the same way as segmental effects such as Final Devoicing and palatalization are found in many unrelated languages. Some sequences of phonological features and segments are found to have an effect on each other when arranged in a certain order. It is natural that these effects take place in the same way in different languages.

In the German sentences examined in this paper, H-raising affects the final accent of an i-phrase. It happens on a narrow focus as well as in half of the realizations of an all-new sentence. As far as narrow focus is concerned, H-raising is completely regular, and this in all positions and in all sentences. If this was the only place where this effect is observed, it would not be possible to distinguish H-raising from an extreme raising of the reference line in the last position of the sentence, due to a narrow focus, and we should conclude that H-raising is always due to narrow focus. However, as mentioned above, H-raising occurs in an all-new context in roughly 50% of the cases. As far as the pragmatic interpretation of an all-new sentence is concerned, all constituents are equally prominent, since they are equally new, and the pitch accents assigned to such a sentence convey this fact. Otherwise, it would be impossible to interpret an all-new sentence correctly. This observation is in line with the assumption of several authors that there is no ‘nuclear’ accent in an all-new sentence (see Selkirk 2002, 2006 for English, and Gussenhoven 1992 for Dutch). In other words, a last accent in a sentence is not semantically or phonologically prominent, even if it is scaled at a higher level than the preceding accents. Nothing in the semantic or in the phonological representation of an all-new sentence, be it a metrical grid, a metrical tree, or any other representation, assigns a special prominence to its last accent. We suggest that the higher level reached by such an accent is due to the H-raising effect just described, a purely tonal effect.

Downstep occurred in the remaining all-new sentences, that is in roughly half of the cases. It may be hypothesized that both H-raising and downstep are available for each speaker and each all-new utterance, but that they conflict as for the effects they exert. The first one is an assimilatory effect of an L on a following H and the second one a dissimilatory effect of an L on the preceding H. Speakers choose between both options each time they have to realize a wide focus sentence, and opt half of the times for downstep and half the times for H-raising.

As far as Final  $f_0$  drop is concerned, an alternative explanation could involve Final Lowering, which has been abundantly described in the literature on intonation languages (see Liberman & Pierrehumbert 1984, Prieto et al. 1996 and Truckenbrodt 2004 among others). In Final Lowering, the final step in a series of downstep is larger than the non-final one. But clearly, Final  $f_0$  drop is a different phenomenon. The final step involved in Final  $f_0$  drop is not the last one of a series of accents, but it is the first of a series of unaccented word, or the only unaccented word, and the drop in  $F_0$  is much larger as what the mentioned authors find in their data for Final Lowering. The only place in our data, where it would make sense to search for Final Lowering is when all constituents are accented, even the verb, since it is the only constellation in which downstep is happening, albeit in 50% of the cases. But we did not look for this effect,

and will not mention it any more (see Truckenbrodt 2004 for a very interesting and elaborate analysis of Final Lowering in German).

As for H-raising, some alternatives are to be considered, as well. First, could it be that H-raising is actually an upstep effect, where upstep is defined as the resetting of an accent to a higher value than normal downstep leads us to expect? Truckenbrodt (2002) as well as Féry & Truckenbrodt (2005) find upstep only in medial phrases, not in sentence-final position. Truckenbrodt (2002) analyzes the German upstep as a return to the phrasal reference-line at the end of the first intonation phrase, and allows the next phrase to be downstepped relatively to this tone. He explains the phenomenon as being triggered by the presence of a following i-phrase. In other words, it is a phenomenon taking place at the boundary between two intonation phrases. The H-raising effect identified in this paper differs from upstep in a number of respects. First, it is located in a final i-phrase, which implies that its role in the tonal phonology is not the same as the phrase-medial upstep effect. It does not provide a key relatively to which a following tone can be downstepped, but rather it allows a final drop to occur. Moreover, it is not necessarily located in the final part of the i-phrase, but can take place relatively early in the sentence, depending of the focus/given structure of the remaining constituents. The third difference between H-raising and upstep as analyzed by Truckenbrodt (2002) is the height of the raised H. The high tone in H-raising does not necessarily return to the value of the first tone, but rather attains a point in the speaker's voice at which it can trigger the final drop. In our data, it is located at roughly 80Hz above the point reached by the final drop. The raised H is, in other words, not scaled relatively to the reference line of the first accent in the sentence, but rather to the last boundary tone.

A second, more functional explanation for H-raising that does not fit our data is provided by Laniran & Clements (2003:232) in the following terms. "[...] although Yoruba speakers implement downstep and H raising by quantitatively different means, their realization strategies "conspire" to insure that downstepping H tone will not penetrate the frequency band reserved for M tones." In other words, H-raising guarantees that a downstepped H is not mistaken for a mid or even a low tone. In German, it is not clear whether there is a tonal domain reserved for low tones. Our speakers do not reset their voice in the same way as Yoruba speakers do.

The last alternative analysis for H-raising is to subsume the tonal effect under consideration to the effects of information structure, which, as we say, was interpreted as raising of the reference line. This alternative does not account for multiple narrow focus, in which all accents are raised by virtue of the raising of the reference line as a consequence of narrow focus, but only the last one is subject to H-raising because it is the only one which precedes the low boundary tone. In other words, the effect of the raising of the reference line and the effect of H-raising are distinct.

### 6.3 H-lowering

Turning now to H-lowering, it was shown in sections 4 and 5, that a high tone immediately preceding a raised high tone is lowered as compared to the same tone preceding a non-raised high tone. This lowering effect is analyzed as the regressive effect of a H\* on a preceding H<sub>P</sub>. The sequence H<sub>P</sub> H\* is target of H-lowering, both in a sequence of tones in an all-new sentence, and in a sequence in which the raised H\* is the narrow focus. As was shown in section 6.2, the sequence L\*H<sub>P</sub> H\* L<sub>i</sub> is the final sequence of tones in a declarative sentence.

For H-lowering, again, two effects are to be distinguished. First the second high tone (the one acting on the preceding tone) is raised. Raising can take place for two reasons: it can be that the reference line is raised because of a narrow focus, or that H-raising has taken place. Second, the first H tone (the one on which the action is exerted) is lowered

and is thus dissimilated from its following tone. For H-lowering, as well, we can lean our analysis on results from the literature on tone languages. Shih (1986), cited in Xu (1997:79) finds that in Chinese, a high tone before another high tone is considerably lower than a high tone before all other tones.

An alternative analysis consists in envisaging the effect as lowering of the reference line. In this case, we predict that the prosodic phrase has lowered its register as a consequence of givenness. That this analysis is not tenable is immediately clear when the results of table 9 are examined. It was shown there that H-lowering only takes place in the high tone immediately preceding the raised H tones. When there are two given elements, it is not the case that both are lowered. This is an indication that the lowering of a high tone only takes place when it precedes a raised high tone. It does not effect all given constituents.

The lower value of a medial tone as compared to its neighbors has been sometimes pointed out in the literature, but until now, none has provided a formal account for it. In our account, H-lowering is a phonological effect which has a phonetic consequence.

The different effects that have been described in the preceding pages interact to produce the tonal contour of a German declarative sentence. The exact  $f_0$  values of tones are calculated from the interaction of several influences. Compositionality is at issue, but it is also true that the  $f_0$  contour is the result of partly conflicting tendencies.

## 7. Conclusion

The present work has examined tonal issues in German with the help of an experiment in which 18 speakers uttered altogether a total of 2277 sentences of the same syntactic structure, but with a variable number of constituents, word order and focus-given structure. This experiment has shown that the scaling of high tones, and thus the entire melodic pattern is influenced by information structure on the one hand, and tonal effects on the other hand. As for the first kind of effect, they are not especially surprising, as German, like English, is an intonation language whose use of accents and accent sequences is conditioned by pragmatic considerations. For the sentences used in the experiment described here, the influence of information structure can be summed up in the following way: Focus raises tones while givenness lowers them in prenuclear position and cancels them postnuclearly. These changes in the values of accents were explained by the influence information structure has on reference lines associated with prosodic domains. However, and crucially, information structure, or normal syntactic structure cannot explain all tonal patterns gathered in the analysis of the 2277 pitch tracks. Further effects are purely tonal ones. First, regular downstep is considered as the default pattern of high tone implementation in a 'neutral' sentence, where neutral is understood as all-new or whole-focused. It can be understood as an effect of downstepping the prosodic reference lines relatively to each other, but we preferred to analyze it as a tonal effect taking place across p-phrases: a  $L^*$  anticipatorily assimilates a following H tone, and thus lowers its value. Second, final  $f_0$  drop is the steep fall from a raised high tone to the bottom line of the speaker and is found in situations of narrow focus, as well as in wide focus, when the last high tone is H-raised. Third, H-raising which has been described among others by Laniran & Clements (2003) for Yoruba, and Xu (1997) for Mandarin Chinese, is shown to be a crucial component of the tonal pattern of German. H-raising is triggered by a low tone exerting a dissimilatory regressive effect on a preceding high tone. We showed that raising a last accent allows the phenomenon of Final  $f_0$  drop to be perceptively prominent. Third, the reverse phenomenon, H-lowering before a raised H is also at play. This effect lowers a high

tone before a raised high tone, and has been analyzed as another dissimilatory regressive effect.

In the discussion, alternative accounts, like upstep, a phenomenon which raises tones at the end of a medial phrase (see Truckenbrodt 2002), and final lowering, the larger final step in a series of downsteps (see Liberman & Pierrehumbert 1984, Truckenbrodt 2004 among others) are shown to be different effects. As far as upstep is concerned, all our phrases were final, since we avoided syntactic complexity, so that the observed pre-final raise cannot be the result of anticipating a following intonation phrase. And clearly, final lowering is not at play in our data, if final lowering is understood as a larger downstep than predicted by the preceding downward steps. In a downstep pattern, the final step was not larger than the preceding ones. The effect of Final  $f_0$  drop cannot be understood as final lowering as it involves a preceding raise on the last accent.

A side-issue of this paper is the existence of pre-planning (or ‘foresight’). It was shown that speakers expand their register when they know that more accents (and more downsteps) are to come. The dissimilatory effects of H-raising and H-lowering can also be considered as proving the existence of a mild form of pre-planning, as a tone has an influence on a preceding one. H-lowering was shown to even involve a kind of double pre-planning.

Our investigation has necessarily been partial since it investigates only sequences of pitch accents in a single intonation phrase. It is evident that the study of more complex intonation patterns will provide a clearer picture of tonal interactions.

## References

- Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of “segmental anchoring” of  $F_0$ : Evidence from German. *Journal of Phonetics*, 32(2), 177–197.
- Barker, G. (2002). Intonation patterns in Tyrolean German: An autosegmental-metrical analysis. Doctoral dissertation, University of California, Berkeley.
- Bartels, C. & Kingston, J. (1994) Salient pitch cues in the perception of contrastive focus. In Bosch, P. & van der Sandt, R. (eds) *Focus and natural language processing*, vol 1, Intonation and Syntax. IBM working papers of the Institute for Logic and Linguistics, 11-28.
- Baumann, Stefan (2006) *The Intonation of Givenness*. Tübingen. Niemeyer.
- Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255–309.
- Berg, Rob van de, Carlos Gussenhoven & Toni Rietveld (1992) Downstep in Dutch: Implications for a model. In Gerhard Docherty & D. Robert Ladd (eds.) *Papers in Laboratory Phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press. 335-367.
- Bierwisch, M. (1966). Regeln für die Intonation deutscher Sätze. *Studia Grammatica*, 7, 99–201.
- Boersma, Paul & Weenink, David (2006). Praat: doing phonetics by computer (Version 4.4.20) [Computer program]. Retrieved May 3, 2006, from <http://www.praat.org/>
- Brown, G., Currie, K & Yule, G. (1980) *Questions of Intonation*. London: Croom Helm.
- Büring, D. (2001). What do definites do that indefinites definitively don't? In C. Féry, & W. Sternefeld (Eds.), *Audiatur vox sapientiae. A Festschrift for Arnim von Stechow* (pp. 70–100). Berlin: Akademie Verlag.
- Cinque, G. (1993). A null theory of phrase and compound stress. *Linguistic Inquiry*, 24, 239–297.
- Clements, G. N. (1990) The status of register in intonation theory: Comments on the papers by Ladd and by Inkelas and Leben. In J. Kingston, & M. Beckman (Eds.),

- Papers in laboratory phonology 1: between the grammar and the physics of speech (pp. 58-71). Cambridge: Cambridge University Press.
- Cohen, A., 't Hart, J., 1967. On the anatomy of intonation. *Lingua* 19, 177-192.
- Cooper & Sorensen. 1983.
- Essen, O. von (1956) Grundzüge der hochdeutschen Satzintonation. Ratingen.
- Féry, C. (1993). German intonational patterns. Tübingen: Niemeyer.
- Féry, Caroline & Vieri Samek-Lodovici (2006) Focus projection and prosodic prominence in nested foci. *Language* 82.1. 131-150.
- Féry, Caroline & Hubert Truckenbrodt (2005) Sisterhood and Tonal Scaling. *Studia Linguistica*. Special Issue "Boundaries in intonational phonology". 59.2/3. 223-243.
- Fitzpatrick-Cole, J. (1999). The Alpine intonation of Bern Swiss German. Paper presented at International congress of phonetic sciences, San Francisco (pp. 941-944).
- Fuchs, A. (1976) 'Normaler' und 'kontrastiver' Akzent. *Lingua* 38, 293-312.
- Gandour, J., Potisuk, S. & Dechongkit, S. (1994) Tonal coarticulation in Thai, *Journal of Phonetics*, 22, 477 - 492
- Grabe, E. (1998). Comparative intonational phonology: English and German. Doctoral dissertation, Universiteit Nijmegen.
- Grice, M., & Baumann, S. (2002). Deutsche Intonation und GToBI. *Linguistische Berichte*, 191, 267-298.
- Grice, Martine, D. Robert Ladd & Amalia Arvaniti (2000) On the Place of Phrase Accents in Intonational Phonology. *Phonology* 17. 143-185.
- Gussenhoven, Carlos (1992) Sentence accents and argument structure. In Iggy Roca (Ed.) *Thematic Structure. Its role in grammar*. Foris. Berlin. 79-106.
- Gussenhoven, Carlos (2002) Intonation and interpretation: phonetics and phonology. In Bel, B. & Marlien, I (eds) *Proceedings of Speech Prosody, Aix-en-Provence: Laboratoire Parole et Language, Université de Provence*. 2002. pp 47-57
- Gussenhoven, Carlos (2004) *The Phonology of Tone and Intonation*. Cambridge University Press.
- Ishihara, Shinichiro (2003) Intonation and interface conditions. Doctoral Dissertation, Massachusetts Institute of Technology.
- Ishihara, Shinichiro (2006) Major Phrase as Spell-Out Domain: Evidence from Japanese Scrambling and NPI. Ms., Potsdam.
- Ishihara, S., Féry, C. (2006). Phonetic Correlates of Second Occurrence Focus. In *Proceedings of the 36th North Eastern Linguistic Society (NESL36)*.
- Jacobs, Joachim (1993) Integration. In: M. Reis (Ed.) *Wortstellung und Informationsstruktur. (Linguistische Arbeiten 306)*. Tübingen. Niemeyer.
- Jun, Sun-Ah, Ed. (2005) *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford University Press.
- Katz, Jonah and Elisabeth Selkirk (2005) Two Types of Focus, Two Types of Prominence. Manuscript, University of Massachusetts Amherst.
- Kügler, Frank (2005) Swabian and Uppers Saxon Intonational Patterns. Dissertation, Potsdam University.
- Kügler, Frank, Féry, Caroline & van de Vijver (2003) Pitch accent realization in German. *Proceedings of the 15th ICPHS, Barcelona, Spain*, 1261-1264.
- Ladd, D. R. (1980) *The structure of intonational meaning: evidence from English*. Bloomington: Indiana University Press.
- Ladd, D. R. (1984). Declination: a review and some hypotheses. *Phonology Yearbook*, 1, 53-74.
- Ladd, D. R. (1990). Metrical representation of pitch register. In J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology 1: between the grammar and the physics of speech* (pp. 35-57). Cambridge: Cambridge University Press.

- Ladd, D. R. (1993) Peak Features and Overall Slope. In Cutler, A. & Ladd, D. R. (eds) *Prosody: Models and Measurement*. Berlin: Springer, 39-52.
- Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Laniran, Y. O. (1992). *Intonation in tone languages: the phonetic implementation of tones in Yoruba*. Ph.D. thesis, Cornell University.
- Laniran, Yetunde O. & G.N. Clements (2003) Downstep and high raising: interacting factors in Yoruba tone production. *Journal of Phonetics* 31. 203-250.
- Lieberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M.Aronoff, & R. T. Oehrle (Eds.), *Language, sound, structure: Studies in phonology presented to Morris Halle by his teacher and students* (pp. 157–223). Cambridge, MA: MIT Press.
- Lieberman, P., 1967. *Intonation, Perception and Language*. Cambridge, MA: MIT Press.
- Maeda, S. (1976) *A characterization of American English intonation*. Doctoral dissertation, MIT, Cambridge, MA
- Mayer, Jörg (1997) *Intonation und Bedeutung*. Arbeitspapiere des IMS. Lehrstuhl für Phonetik. Universität Stuttgart.
- Peters, Jörg (2005) *Intonatorische Variation im Deutschen*. Studien zu ausgewählten Regionalsprachen. Habilitationsschrift. University of Potsdam.
- Pheby, J. (1980): *Phonologie: Intonation*. In Heidolph, K. E., W. Flämig, W. Motsch u.a. *Grundzüge einer deutschen Grammatik*, Chapter 6, 839-897. Berlin: Akademie-Verlag.
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Doctoral dissertation, MIT.
- Prieto, P., D'Imperio, M., Elordieta, G., Frota, S., & Vigário, M (2006). Evidence for *soft* preplanning in tonal production: Initial scaling in Romance. *Proceedings of Speech Prosody*, Dresden.
- Prieto, P., Shih, C., & Nibert, H. (1996). Pitch downtrends in Spanish. *Journal of Phonetics*, 24, 445–473.
- Rialland, A., & Somé, P. A. (2000). Dagara downstep: How speakers get started. In V. Carstens, & F. Parkinson (Eds.), *Advances in African linguistics. Trends in African Linguistics*, Vol. 4 (pp. 251–263). Trenton, NJ: Africa World Press.
- Selkirk, Elisabeth (2002) Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in English. In *Speech Prosody 2002: Proceedings of the 1st International Conference on Speech Prosody*, Aix-en-Provence, 643-646.
- Selkirk, Elisabeth (2006) The Prosodic Phrasing of FOCUS followed by focus: The Case of English. In Gussenhoven, Carlos and Tomas Riad (Eds.) *Tones and tunes: studies in word and sentence prosody*. Berlin: Mouton.
- Stewart, J. M. (1993). Dschang and Ebr!e as Akan-type total downstep languages. In H. van der Hulst, & K. Snider (Eds.), *The phonology of tone: The representation of tonal register* (pp. 185–244). Berlin: Mouton de Gruyter.
- Thorsen, N. (1983). Standard Danish sentence intonation: Phonetic data and their representation. *Folia Linguistica*, 17, 187–220.
- Truckenbrodt, Hubert (2002) Upstep and embedded register levels. *Phonology* 19. 77-120.
- Truckenbrodt, Hubert (2004) Final lowering in non-final position. *Journal of Phonetics* 32. 313-348.
- Uhmann, S. (1991). *Fokusphonologie*. Tübingen: Niemeyer.
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61–83.

[f.kuegler@googlemail.com](mailto:f.kuegler@googlemail.com)