Available online at www.sciencedirect.com
ScienceDirect
Journal of Phonetics 36 (2008) 680–703
Journal of Phonetics
www.elsevier.com/locate/phonetics

Pitch accent scaling on given, new and focused constituents in German

Caroline Féry*, Frank Kügler

Department of Linguistics, University of Potsdam, Karl-Liebknecht-Straße 24/25, D - 14476 Golm, Germany

Received 20 October 2006; received in revised form 28 April 2008; accepted 6 May 2008

Abstract

The influence of information structure on tonal scaling in German is examined experimentally. Eighteen speakers uttered a total of 2277 sentences of the same syntactic structure, but with a varying number of constituents, word order and focus-given structure. The quantified results for German support findings for other Germanic languages that the scaling of high tones, and thus the entire melodic pattern, is influenced by information structure. Narrow focus raised the high tones of pitch accents, while givenness lowered them in prenuclear position and canceled them out postnuclearly. The effects of focus and givenness are calculated against all-new sentences as a baseline, which we expected to be characterized by downstep, a significantly lower scaling of high tones as compared to declination. The results further show that information structure alone cannot account for all variations. We therefore assume that dissimilatory tonal effects play a crucial role in the tonal scaling of German. The effects consist of final f0 drop, a steep fall from a raised high tone to the bottom line of the speaker, H-raising before a low tone, and H-lowering before a raised high tone. No correlation between word order and tone scaling could be established.

1. Introduction

Researchers agree on classifying German as an intonation language (see the typologies in Gussenhoven, 2004; Jun, 2005; Ladd, 1996, among others). It has lexical stress and uses different types of postlexical pitch accents and pitch accent sequences as well as boundary tones for the expression of pragmatic contrasts (Baumann, 2006; von Essen, 1956; Féry, 1993; Mayer, 1997; Peters, 2006; Pheby, 1980). Various aspects of German dialect intonation are well studied (Barker, 2002; Gilles, 2005; Kügler, 2007; Peters, 2006), but here we restrict ourselves to Standard German as spoken in the Berlin-Brandenburg region. Thus far, the bulk of the analysis of the tonal structure of this language has been directed towards the tone inventory, the alignment of these tones with segments, and variation in the use of tones. Tone scaling across accents and across phrases has not been extensively studied, though some research has been done on the subject (see Katz & Selkirk, 2005; Ladd, 1990 and Selkirk, 2006 for English; Féry & Truckenbrodt, 2005; Kügler, Féry, & van de Vijver, 2003; Truckenbrodt, 2002 for German). A number of studies on German have specifically investigated

*Corresponding author. Tel.: +49 331 977 2432; fax: +49 331 977 2087.
E-mail addresses: caroline.fery@googlemail.com (C. Féry), f.kuegler@googlemail.com (F. Kügler).

0095-4470/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2008.05.001

downstep (Grabe, 1998; Truckenbrodt, 2002, 2004).[1] To our knowledge, however, no study has been published so far that addresses tone scaling in German in relationship with information structure more than in passing, and purely tonal effects like H-raising and H-lowering have been left unnoticed. However, to fully understand the tonal structure of a language such as German, a quantification of the combined effects of syntax, information structure and of the tones themselves is necessary. This is the subject of the present paper.

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 appeared in two variants: wide (or all-new) and narrow (see Ladd, 1980 for this distinction). In an all-new sentence, no element has been mentioned in the preceding context or was especially prominent in the common ground of the protagonists. A narrow focus was induced by a context asking explicitly for one or more arguments, or for the verb. By contrast, a given constituent has been mentioned in the question or the context introducing the target sentence. A semantically informed definition of focus and givenness is beyond the scope of this paper, but we base our view of information structure on the work of Rooth (1985, 1992), Schwarzschild (1999), and the papers in Féry, Fanselow, and Krifka (2006), among others.

Default sentence accent assignment, as well as accents motivated by information structure, have been examined extensively for German (for instance by Bierwisch, 1966; Büring, 2001; Cinque, 1993; Féry & Samek-Lodovici, 2006; Fuchs, 1976; Gussenhoven, 1992; Jacobs, 1993; von Stechow & Uhmann, 1984). In an all-new sentence, accents are assigned on metrically prominent positions in the sentence, determined solely 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; Gussenhoven, 1992; Jacobs, 1993). In the all-new pattern, all arguments are expected to be accented. According to Pierrehumbert (1980) and Liberman and Pierrehumbert (1984), scaling of tones in English takes place from left to right, downstepping a high tone relative 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.

At the end of this subsection, a series of hypotheses are formulated, that summarize our expectations concerning the experimental results. In German, as well, downstep of a series of accents has been taken to be the unmarked pattern (see, for instance, Féry, 1993; Truckenbrodt, 2002, 2004). The all-new sentence condition can thus function as a baseline for comparison of the effects of information structure (Hypothesis 1).

In sentences with narrow foci, main accent is assigned to the last focused constituent. Based on the findings for different Germanic languages that focus generally boosts accents (Baumann, Grice, & Steindamm, 2006; Cooper, Eady, & Mueller, 1985; Eady & Cooper, 1986; Eady, Cooper, Klouda, Mueller, & Lotts, 1986; Grønnum, 1992; Thorsen, 1985) we expect that a narrow focus raises a pitch accent in all positions, as formulated in Hypothesis 2. In the same way, as focus is expected to raise the values of pitch accents, givenness is expected to lower them (Bartels & Kingston, 1994; Brown, Currie, & Kenworthy, 1980; Cruttenden, 2006; Féry & Ishihara, 2005, to appear; Ladd, 1980, 1983), though prenuclear and postnuclear givenness have completely different effects. We expect given constituents to be realized lower in prenuclear position and to be deaccented in postnuclear position (Hypothesis 3). A prenuclear given constituent is associated with 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 analyzed as being deaccented (see also Cruttenden, 2006; Ladd, 1996; Liberman & Pierrehumbert, 1984 for English).

In Fig. 1, the expectations related to the effect of narrow focus and givenness at different places in a sentence with multiple arguments and a final verb are displayed. The dotted line shows the regular downstep in an all-new sentence. This is the pattern assumed to be the unmarked intonation according to Hypothesis 1. Fig. 1a stands for an initial narrow focus on argument one (A1). A1 and A2 and stand for Argument 1 and 2,

[1]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 or phrasal boundary line 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.

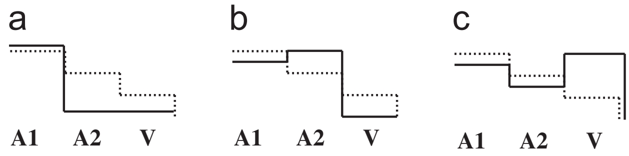

Fig. 1. Expected effect of narrow focus (solid line) in comparison to the regular downstep pattern of the all-new sentence condition (dotted line): (a) focus on an initial argument (A1), (b) focus on a non-initial argument (A2), and (c) focus on a final verb (V).

respectively. A1 is higher than in the all-new configuration (because it is narrowly focused, Hypothesis 2), argument 2 (A2) and the verb (V) are lower (because they are given and deaccented). Fig. 1b 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, Hypothesis 3), A2 is expected to be higher (it is narrowly focused), and V lower. Fig. 1c shows the changes in pitch accents when the verb is narrowly focused. The preverbal arguments are lower than in the all-new condition and downstepped relative to each other, since both are given, and there is an upstep on the verb.

In addition to the effects on f0 we also expect an effect of information structure on duration (Hypothesis 4). Durational influences of focus have been shown for English (Beaver, Clark, Flemming, Jaeger, & Wolters, 2007; Eady & Cooper, 1986), for German (Baumann et al., 2006; Kügler, 2008) and in the case of second occurrence focus in German (Féry & Ishihara, 2005, to appear). We therefore expect a lengthening effect of focus to occur in our data as well.

The last, more speculative, hypothesis for information structure relates to word order. It is assumed that non-canonical word orders, that is those in which an argument other than the nominative is initial, are always triggered by marked information structure, and that, for this reason, scaling of pitch accents should be affected by changes in word order (see Hypothesis 5).

Following Liberman and Pierrehumbert (1984) for English and Grabe (1998) for German we predict a final constant value at the end of declarative sentences ('low endpoint' in the terminology of Liberman & Pierrehumbert), regardless of the length of the sentence or of its focus pattern (see Maeda (1976) for introducing this notion and Menn and Boyce (1982) for discussion, Hypothesis 6).

Finally, based on the discussion of information structure and sentence accent assignment above we expect that the combined effects of syntax and information structure explain all tone values found in the experiment (Hypothesis 7).

The hypotheses are formulated as follows:

1. All-new sentences induce an unmarked accent pattern in which all constituents carry an accent, and there is a downstepping pattern throughout; the verb is optionally accented.
2. Narrow focus on a constituent induces raising of the corresponding high tone.
3. Prefocal givenness induces lowering of the high tone and postfocal givenness induces final deaccenting.
4. A constituent with narrow focus is longer than the corresponding constituent in an all-new sentence or when it is given.
5. A changed word order increases the pitch height of the prominent constituent.
6. A final constant value ('low endpoint') is reached at the end of declarative sentences, regardless of the length of the sentence or of its information structure.
7. The pitch height of all individual pitch accents are unambiguously determined by default prosodic phrasing and information structure.

The next section outlines the methodology of the production experiment. Section 3 reports the results and quantifies the effects of newness, givenness and narrow focus on the scaling of high tones in pitch accents in German. The results of the production experiment revealed that information structure alone cannot account for the various patterns in our data. Therefore, we propose additional dissimilatory tonal effects that account well for the additional effects. In Section 4, a model of pitch accent scaling is proposed in terms of register reference lines that integrate the effects of syntax and information structure, as well as the purely tonal effects introduced in Section 3.

2. Method

2.1. 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 to 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.

The reason for choosing an embedded verb-final pattern is that this structure is considered to be the most unmarked word order of German (Lenerz, 1977). 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 an accusative. In the experiment, all sentences were introduced by a context, which was just a single question (in the case of all-new sentences) 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, but not the masculine). The five animal names used were trochaic with a final schwa syllable (*Hammel* 'sheep,' *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 were located in 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 (1)–(4), embedded in the appropriate context. The material between 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.

(1) **ND**V (focus on two arguments but not 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 sheep the buck followed has
'Because the sheep followed the buck.'

(2) **NAV** (all-new condition)
{Why were the animals happy?}
Weil der Hammel den Rammler eingeladen hat.
'Because the sheep invited the buck.'

(3) NADV (focus on the dative complement)
{The sheep wanted to introduce the buck to the lion. Why didn't he do this?}
Weil der Hammel den Rammler dem Hummer vorgestellt hat.
'Because the sheep introduced the buck to the lobster.'

(4) NV (focus on the verb)
{The animals don't like to fight. Why are they angry with the lobster?}
Weil der Hummer angefangen hat.
'Because the lobster started (a fight).'

Three parameters, which are listed in Table 1, were systematically varied. These parameters comprise the number of arguments, word order and information structure. Altogether 26 conditions were created.[2] Each condition was realized in five versions, in which the five animal nouns were permuted in a systematic way, but the verb remained constant. In total, this resulted in 130 sentences (26 conditions $\times$ 5 renditions) per speaker. In total 2340 sentences (130 sentences $\times$ 18 speakers) were recorded.

[2]The missing constellations comprise very unnatural possibilities, like focus on an initial scrambled constituent.

Table 1
Parameters varied in the production experiment; focused constituents are underlined; altogether 26 conditions

All-new	Narrow focus on one argument	Narrow focus on the verb	Focus on all arguments
NV	<u>N</u>V	N<u>V</u>	
NAV	<u>N</u>AV	NA<u>V</u>	<u>NA</u>V
	N<u>A</u>V		
	A<u>N</u>V		
NDV	<u>N</u>DV	ND<u>V</u>	<u>ND</u>V
	N<u>D</u>V		
	D<u>N</u>V		
NDAV	<u>N</u>DAV	NDA<u>V</u>	<u>NDA</u>V
	N<u>D</u>AV		
	NA<u>D</u>V		
	AD<u>N</u>V		
	AN<u>D</u>V		
	DA<u>N</u>V		
	ND<u>A</u>V		
	DN<u>A</u>V		

2.2. Recordings

The entirety of the data used in this paper was 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 20s, coming from the Northern area of Germany. They were paid for their time.

The target sentences were recorded in a soundproof booth on a DAT tape recorder. A set of instructions familiarized the subjects with the process and had 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 previously. They were spoken by the second author, a trained phonetician and native speaker of Standard German in his thirties, who also comes from the northern part of Germany. He was recorded in a soundproof booth on a DAT recorder. He spoke naturally, at a conversational tempo. The target sentences of the present experiment were intermingled with filler sentences from other experiments.

2.3. Analysis

Of the total of 2340 sentences uttered by the 18 speakers (130 each) 63 were not considered in the final results because of non-measurable accents, mostly due to creaky voice. Altogether 2277 sentences were retained for analysis. The recorded sentences were digitized at a sampling rate of 16 kHz. 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 due to the f0 algorithm (for instance octave jumps) were corrected by hand, and the contour was smoothed using the Praat smoothing algorithm (frequency band 10 Hz) to minimize microprosodic perturbations. All frequency measurements were semi-automatically done using a script that detects the highest f0 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 (5) with '#' as the indicator of a boundary.

(5) # Weil # der Reiher # den Hummer # eingeladen # hat #
 # because # the.NOM heron # the.ACC lobster # invited # has #

Table 2
Distribution of pitch accents on the different verbs in the all-new sentence conditions in percent; numbers of tokens are given in parentheses

	Accented verb	Unaccented verb
NDAV (*vorgestellt*)	50% (45)	50% (45)
NDV (*nachgelaufen*)	75% (64)	25% (21)
NAV (*eingeladen*)	73% (61)	27% (22)
NV (*angefangen*)	86% (77)	14% (13)

The analysis was done in a number of steps. Firstly, a Praat script located f0 maxima in each of the domains, thus in (5), five f0 maxima were identified. Secondly, the result of the Praat script was hand-edited to correct spurious labeling. The authors inspected the tone labels against the f0-track, the substring divisions, an auditory impression, and the spectrogram. Where the more flexible Praat script had assigned an H label that was not in the position where the more narrow criteria above would place it (because of obvious errors due to the algorithm), the label was manually moved. Thirdly, another Praat script recovered the f0-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, as shown in (5).

For statistical analyses, the obtained Hz values were aggregated within each participant and each condition. These aggregated scores were subjected to a repeated measures ANOVA in the case of focus and givenness, and to paired-sample *t*-tests in case of tonal effects, with speakers as random factor. The figures and tables show the aggregated scores per condition averaged across speakers.

3. Results

3.1. All-new sentences

According to our Hypothesis 1 we expected that in the case of wide focus, all arguments would be accented. Depending on the extent of integration with the preceding argument, we further expected the verb to be sometimes accented and sometimes not. Moreover, we expected a downstepped pattern in all four conditions. The first two expectations were fulfilled while the last one has to be revised.

Altogether 348 realizations of all-new sentences were analyzed in the four all-new sentence conditions. All arguments were associated with a pitch accent in all instances of an all-new sentence. 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).[3] The ratio between the groups *accented* and *unaccented* verb across sentence conditions is illustrated in Table 2. As can be seen from these data, the verb was accented more often in a short sentence than in a longer one. All speakers except one (speaker 6) showed free variation in the accenting on the verb (see Appendix A for an overview of the speaker variation).

In (7) and Fig. 2, both versions of a wide focused sentence are illustrated with the same speaker. In (7a), the verb is accented, whereas it is unaccented in (7b). Recall that a sentence that 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 context. It differs from a sentence with a narrowly focused constituent in lacking a given (previously mentioned) constituent. From the representative pitch tracks in Fig. 2, it is clearly visible that the verb has a pitch accent in the first case, but not in the second.

(7) Why were the animals happy?
a. (Weil der HUMmer dem Löwen den RAMMler VORgestellt hat)$_F$
'Because the lobster introduced the buck to the lion.'

[3]The large number of accented verbs in the all-new 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 verbs would be left unaccented.

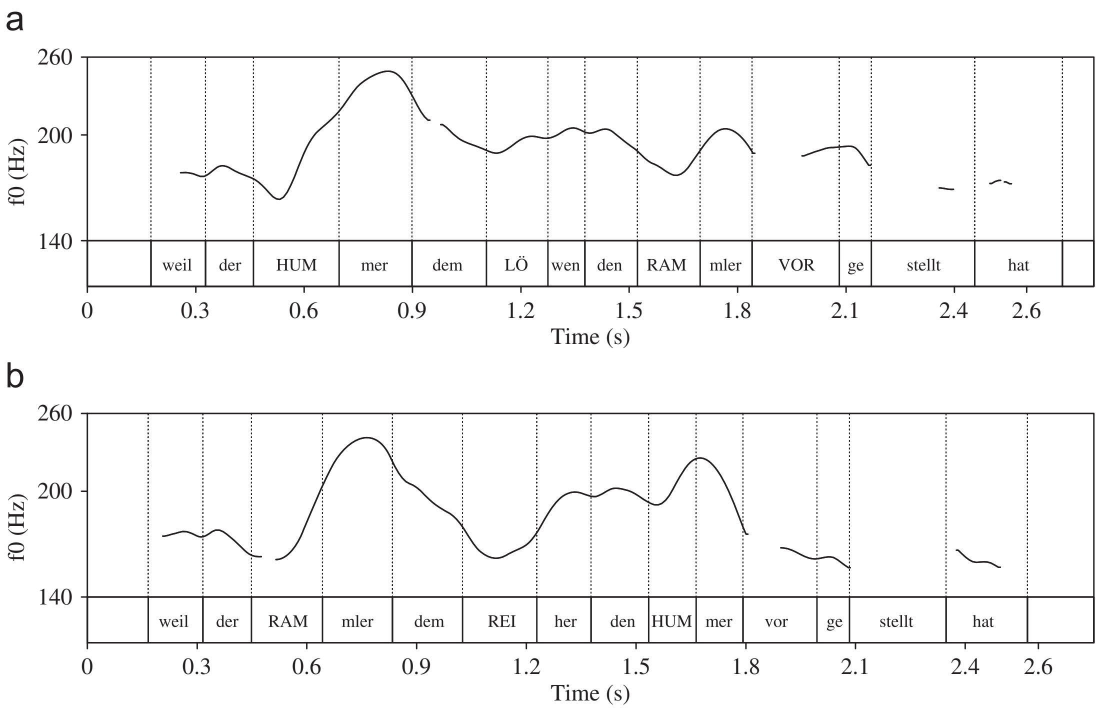

Fig. 2. Two pitch tracks of all-new sentences with three arguments (NDAV): a. shows an accented verb. b. shows an unaccented verb.

b. (Weil der RAMMler dem REIher den HUMmer vorgestellt hat)$_F$
Because the buck introduced the lobster to the heron.'

Turning now to accent scaling, the expectation that the pattern would show downstep everywhere was not met.

Two large classes of cases must be distinguished: In 45.7% of the sentences, a regular downstep took place, in which each accent was lower than the preceding one. But in the remaining 54.3%, raising of an accent was observed. The raised constituent was either on the verb, in which case, of course, the verb was accented (Fig. 2a), or on the preverbal argument, in which case the verb was unaccented (see Fig. 2b).[4,5] In a pattern involving raising, the last accent interrupted downstep and was realized much higher than predicted by a regularly descending pattern.

Not a single speaker used the downstep pattern in every all-new sentence, but all speakers used it at least once. The same is true for the upstep pattern: Every speaker produced at least one utterance with an upstep on a constituent, but nobody used the upstep pattern consistently (see Appendix A).

Fig. 3 shows the average f0 value for each accent in all patterns across all speakers. First, when the verb was accented, three patterns were identified: downstep throughout, upstep on the preverbal argument, upstep on the verb. Second, when the verb was unaccented, two patterns were realized: downstep throughout, and upstep on the preverbal argument. The values in the cells and on the graphs, if nothing indicates the contrary, are averaged high tones for all speakers and all utterances.

Some values were constant. The unaccented verbs had an average f0 varying between 186 and 190 Hz, and the pitch of the raised argument preceding the unaccented verb was also constant, between 267 and 280 Hz.

[4]In a few cases (15 realizations altogether), the raising took place on the final argument preceding an accented verb.

[5]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 did not place 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.

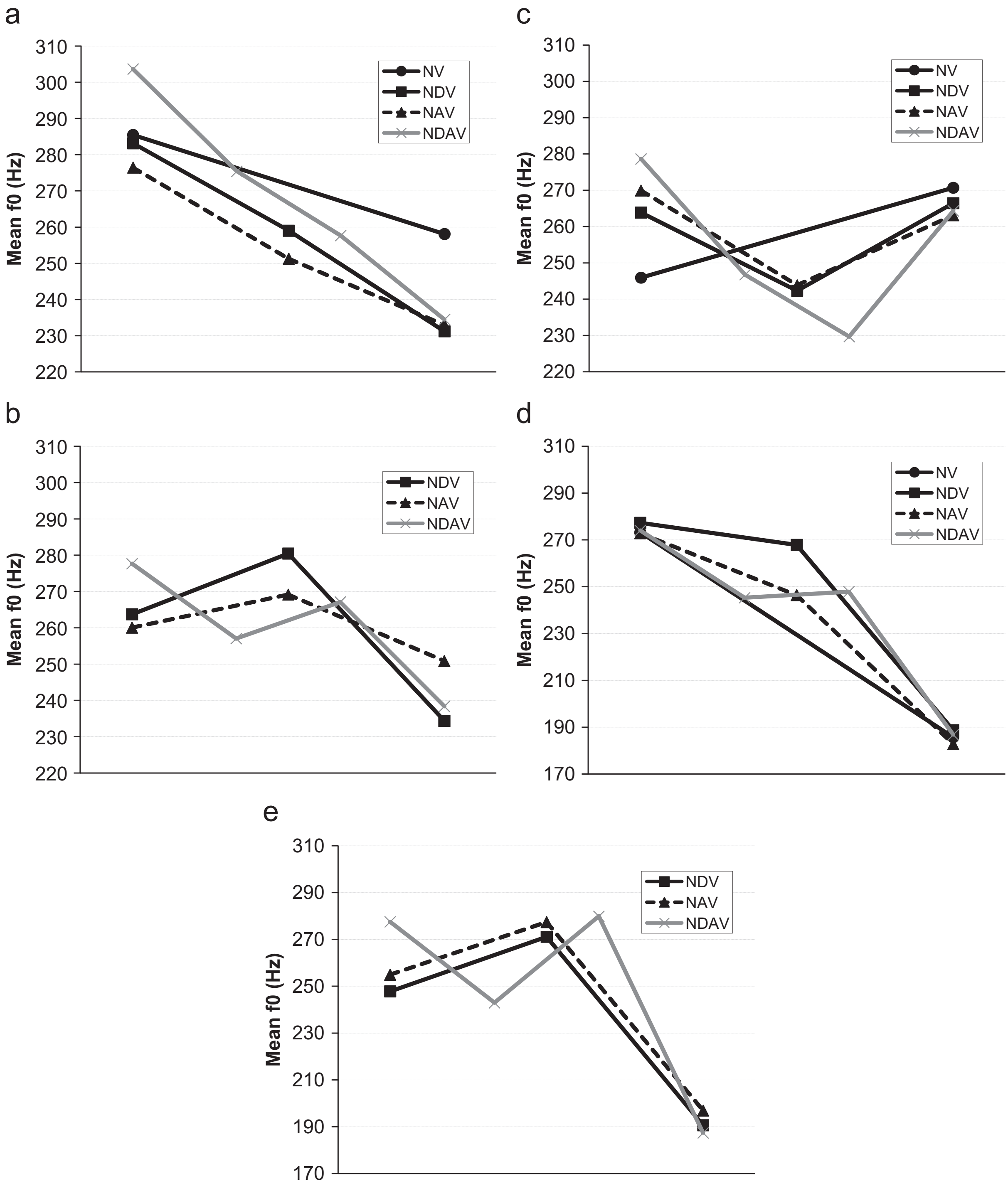

Fig. 3. Mean f0 of all-new sentences in different realizations: (a–c) accented verb, (d–e) unaccented verb; the percent of occurrences relates to the total number of 348 all-new sentence realizations. (a) regular downstep, accented verb; 33.9%. (b) upstepped argument, accented verb; 4.3%. (c) upstepped verb, accented verb; 32.8%. (d) regular downstep, unaccented verb; 11.8%. (e) upstepped argument, unaccented verb; 17.2%.

3.2. Effect of narrow focus on pitch accents

A narrow focus systematically induced a bitonal pitch accent on the narrowly focused argument or verb. Hypothesis 2, which posited that a pitch accent is raised in a narrow focus configuration relative to its value in an all-new sentence, was entirely confirmed.

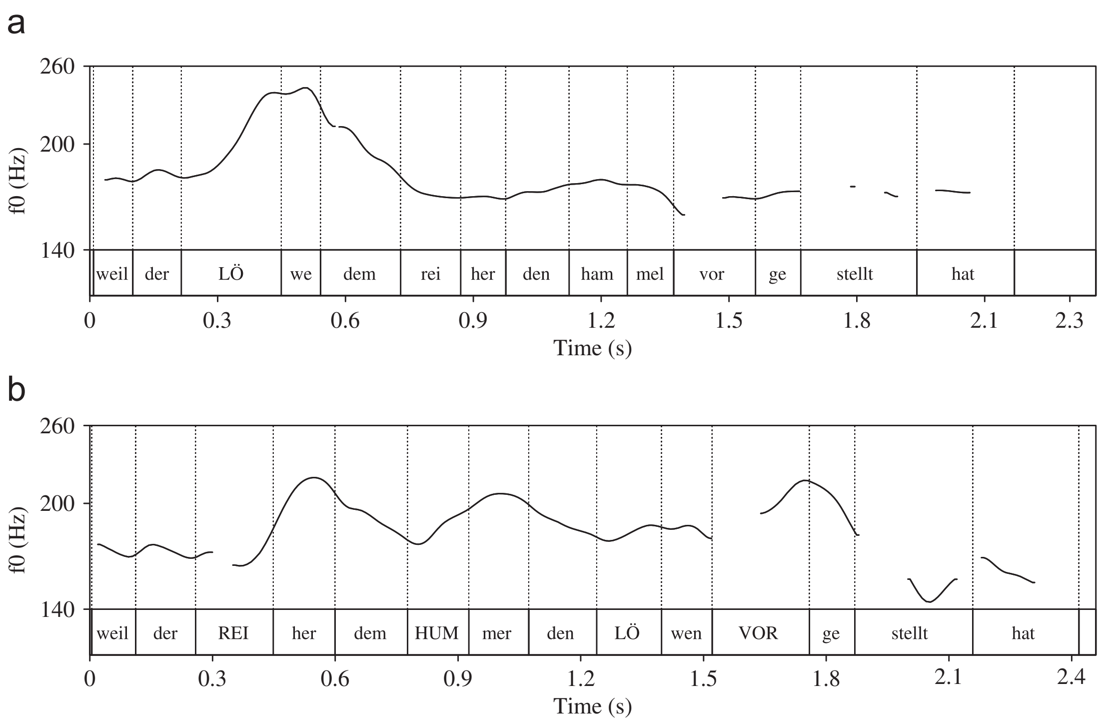

Fig. 4. Pitch tracks of sentences with narrow focus. a. shows narrow focus on the initial constituent. b. shows narrow focus on the verb.

Narrowly focused sentences are illustrated in (8) and in Fig. 4. Fig. 4a shows a narrow focus on the nominative, in which case the first argument is accented and all following arguments and the verb are unaccented (8a). Fig. 4b illustrates a narrow focus on the final verb (8b). In this case, all preceding arguments have a pitch accent. In other words, the postnuclear deaccented material in Fig. 4a is much flatter than the unfocused prenuclear material in Fig. 4b.

(8) a. Weil (der Löwe)$_F$ dem Reiher den Hammel vorgestellt hat.
'Because the lion introduced the sheep to the heron.'
b. Weil der REIher dem HUMmer den Löwen (VORgestellt)$_F$ hat.
'Because the heron introduced the lobster to the lion.'

Fig. 5 shows the mean pitch for the sentences with a unique narrow focus (a: A1, b: A2, c: A3, d: V). A pitch accent on a narrow focus was higher than the corresponding accent in an all-new sentence with downstep, as can be gathered from a comparison with the values for A1, A2, and A3 in all-new conditions (the cases that show the unmarked downstep pattern) from Table 3.[6] A three-way repeated measures ANOVA with the factors *focus*, *position of the arguments*, and *number of arguments* confirmed the difference between all-new and narrow focus sentences with a main effect of *focus* ($F(1,17) = 15.85$; $p < .001$). A significant interaction of *focus* and *position* ($F(1,17) = 11.50$; $p < .005$ for position one and two; $F(1,17) = 12.93$; $p < .005$ for position two and three) showed that this difference depends on the position of the argument, the effect being stronger for A2 than for A1 and A3. The interaction of *number of arguments* and *focus* showed no significant effect for one- and two-argument structures. Yet a significant difference was found between two- and three-argument structures ($F(1,17) = 9.51$; $p < .01$), meaning that the scaling of longer sentences differed significantly among focus structures. Finally, the interaction of *focus*, *number of arguments,* and *position* taken together showed a significant effect ($F(1,17) = 10.34$; $p < .005$), which indicated that the two factors *position of the arguments* and *number of arguments* had an impact on the scaling of the accents in different focus structures.

[6]This is not true of the value of the nominative in a NV sentence. The reason is that only the all-new sentences realized with a downstep pattern entered this table. If the results of the sentences with upstep on the verb displayed in Fig. 3c were taken into account as well, the pattern would be the same as in the other conditions.

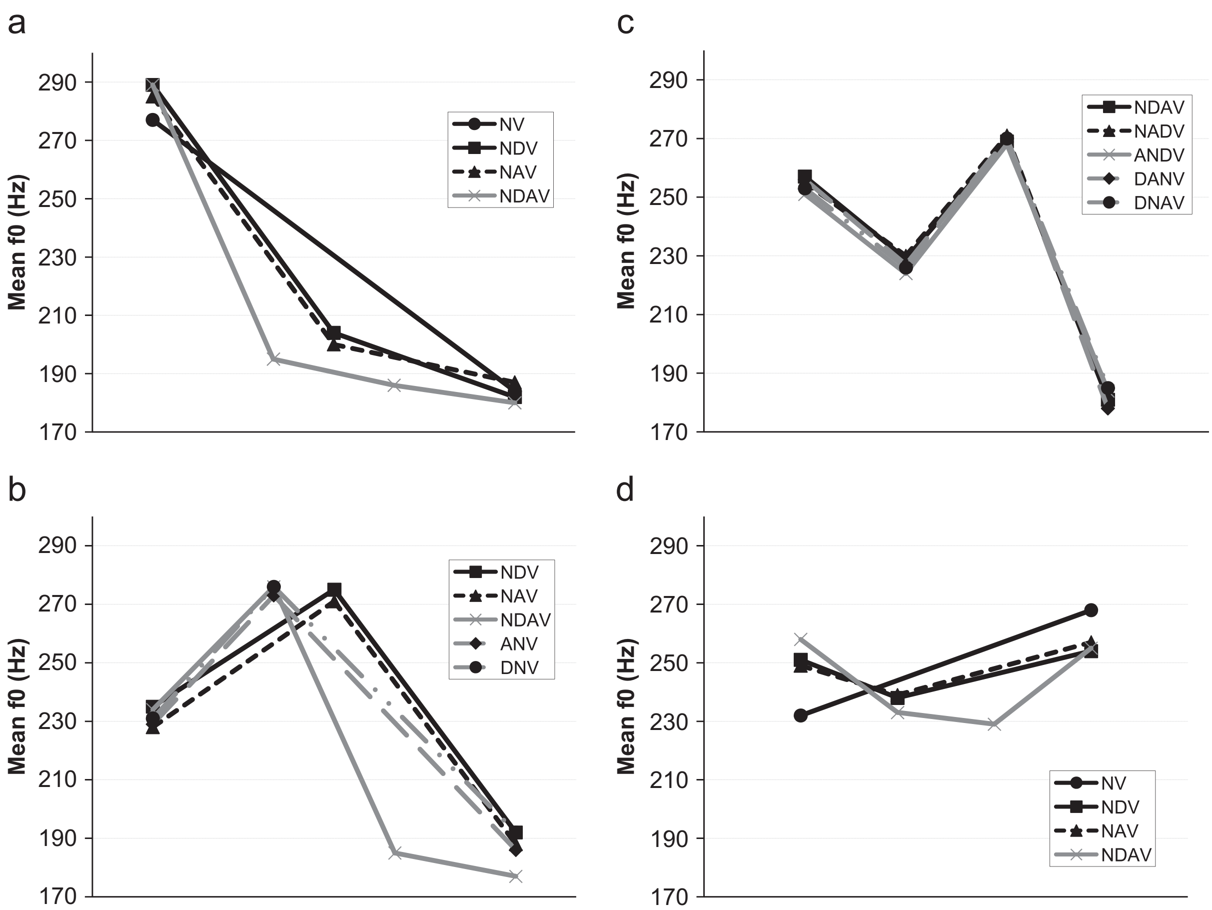

Fig. 5. Mean f0 of narrow focus on different constituents. (a) Narrow focus on the first argument. (b) Narrow focus on the second argument. (c) Narrow focus on the third argument. (d) Narrow focus on the verb.

Table 3
Comparison of the high tone (in Hz) of the arguments (A1–A3 and the verb V) in the downstep pattern of all-new sentences with that of the corresponding arguments in narrow focus

	All-new sentence (downstep)				Focus on the underlined constituent			
	A1	A2	A3	V	A1	A2	A3	V
NDV	282	261		222	289	275		254
NAV	276	250		224	285	271		257
NDAV	293	265	254	217	289	276	269	255
NV	282			238	277			268

The remaining patterns are those in which more than one argument is focused, but the verb is not; see the values of the high tones in Table 4. In conditions NDV and NAV, two arguments are focused (A1 and A2), and in condition NDAV, all three arguments are focused. In the latter configuration, downstep applied between the first two arguments. The difference in height between A1 (275 Hz) and A2 (243 Hz) was on average and across all speakers 32 Hz, and between A2 and A3, 20 Hz. Upstep applied on A3. In the sentences with two arguments, the upstep occurred on the second argument, which was immediately preverbal, and the effect of downstep was obscured or cancelled out.

An interesting observation, shown in Table 3 above, is that an initial narrowly focused A1 was higher in pitch than a narrowly focused A2, which in turn was higher than a narrowly focused A3. More generally, constituents decreased in pitch height from the beginning to the end of the sentence across conditions. This can

Table 4
Pitch peak values in Hz on the arguments and the verb in cases where all arguments are focused, but the verb is not; focused constituents are underlined

	A1	A2	A3	V
NDV	264	270		188
NAV	257	269		189
NDAV	275	243	263	183

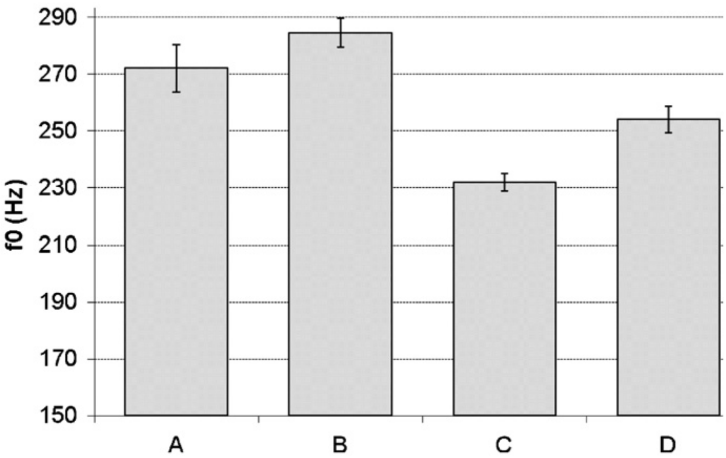

Fig. 6. Mean f0 showing a 95% confidence interval for the initial nominative in different environments: (A) Nominative in an all-new pattern. (B) Narrowly focused nominative. (C) Given nominative before a narrow focus. (D) Given nominative before a given argument.

be interpreted as an effect similar to downstep. The narrowly focused constituents later in the sentences were preceded by prefocally accented constituents.

Finally, in line with all other constituents boosted by a narrow focus, a focused verb had a higher f0 when it was narrowly focused than when it was part of an all-new sentence or when it was unaccented (see Fig. 5d). However, f0 on an accented verb was definitely lower than f0 on a narrowly focused argument. The narrowly focused verb was only slightly higher or even lower than A1, except for in the NV pattern, in which the verb was much higher than the unique argument.

To sum up this section, a narrow focus had the effect of raising an accent's f0 locally. A narrowly focused word's f0 was always significantly higher than the same word in a downstepped pattern of an all-new sentence.

3.3. *Effect of givenness on pitch accents*

3.3.1. *Prefocal givenness*

Hypothesis 3 was also confirmed. Prefocal givenness lowered f0 of pitch accents, but did not delete them. Perceptually, and also compared with postfocal deaccented constituents, prefocal arguments had pitch accents, although these were lower than their equivalents in a focused context, as is shown in Fig. 6. The f0 of an initial given nominative (the bars C and D in Fig. 6) was lower than an initial nominative in an all-new sentence (A) and in a narrow focus (B). These results show a difference between the two prefocal givenness conditions: f0 of the given nominative was lower prior to a narrow focus than prior to another given constituent. This difference cannot be explained by means of information structure alone, and we come back to these data in Section 3.7 below.

Also sentence-medially, prefocal lowering applies. When A2 was given, its f0 varied on average across all speakers between 224 and 230 Hz. When it was part of an all-new pattern, its f0 was on average between 253 and 257 Hz. And when A2 was narrowly focused, the f0 range was between 271 and 276 Hz.

Table 5
Pitch peak values in Hz in initial given arguments (A1, A2, and A3) preceding a narrow focus (underlined constituent in the respective condition)

Condition	A1	A2	A3
NAV	249	239	
NDV	251	238	
NDAV	257	229	
NADV	255	230	
ADNV	255	227	
ANDV	251	224	
DANV	256	227	
DNAV	253	226	
NDAV	258	233	229

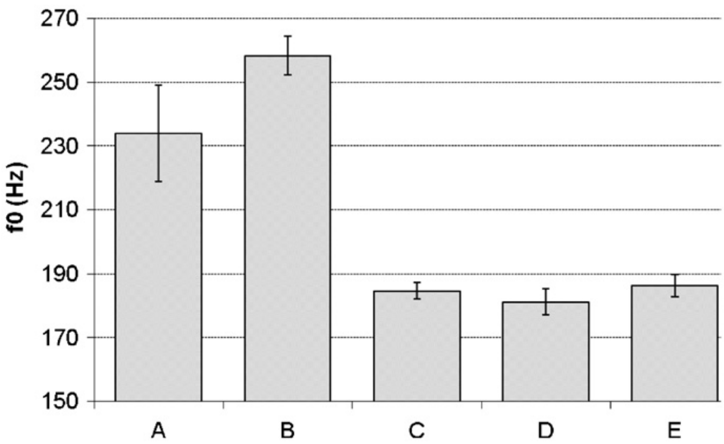

Fig. 7. Mean f0 showing a 95% confidence interval for the effect of givenness on the verb. (A) All-new condition. (B) Narrow focus on the verb. (C) Narrow focus on preverbal argument. (D) Narrow focus on pre-preverbal argument (preverbal argument and verb are given). (E) All preverbal arguments are focused.

As illustrated in Table 5, a sequence of two prefocal given arguments (A1 and A2) showed downstep. The downward steps were about 12 Hz on average when the two given arguments preceded a narrowly focused verb. When the narrowly focused constituent was an argument, the downward steps were larger: around 27 Hz.

3.3.2. Postfocal givenness

In the postfocal position, arguments were deaccented, in confirmation of Hypothesis 3. First, Fig. 7 compares the value of the verb in different environments. Columns A and B show f0 of the verb in an all-new sentence and in a narrow focus environment, respectively. Columns C to E show the value of a verb in a given context when the arguments have different focus patterns (see legend). It can be observed from these data that, in the postnuclear position (thus C–E), the verb has a very stable value.

The same deaccenting effect could be seen on postnuclear medial values, which were lower when the constituents were given than when they were new. This is shown in Table 6 for some relevant conditions. The initial constituent is always focused, and the medial argument is given in the first/second columns and focused in the third/fourth ones.

Finally, there is evidence that given constituents are in a downstep relationship to each other. When a second (and third) postnuclear constituent was present, downward steps were still realized, even though they became much smaller at the end of the sentence, as Fig. 5a shows and as Fig. 5b suggests. We do not try to

Table 6
Pitch peak values in Hz for medial arguments in a postfocal environment (left column) compared with corresponding narrowly focused arguments; focused constituents are underlined

	Postfocal		Narrowly focused
NAV	200 (A)	NAV	269 (A)
NDV	204 (D)	NDV	270 (D)
NDAV	195 (D)	NDAV	243 (D)
NDAV	185 (A)	NDAV	263 (A)

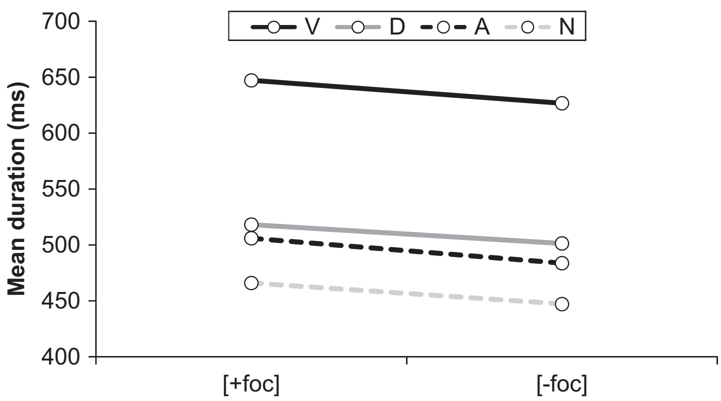

Fig. 8. Effect of focus on duration. An interaction plot of type of focus by syntactic category for the dependent variable mean word duration.

interpret this decrease in pitch in the postnuclear stretch of the sentence in this paper, but pursue this issue in an independent experiment.[7]

3.4. *Effect of narrow focus on duration*

Hypothesis 4 predicts that narrow focus has an effect on duration: a constituent which is narrowly focused is expected to be longer than when it is part of an all-new sentence (see Cooper et al. (1985) for such an effect in English). The hypothesis was confirmed in our data. Narrowly focused constituents were significantly ($F(1,17) = 20.775$; $p < .001$) longer than non-focused ones, a result shown in Fig. 8. A second result is also apparent from this figure: nominative arguments were shorter than their accusative and dative correspondents. This difference may reflect a difference in their position in the sentence since a nominative occurred initially more often than the other constituents, though more data are needed to confirm this hypothesis. The longer duration of the verb comes about because it is quadrisyllabic (except *vor.ge.stellt* 'introduced'), as opposed to the arguments, which are always trisyllabic (article + noun).

A further effect is that duration can be related to the number of constituents that a sentence contains. Fig. 9 shows that the duration of an initial nominative increased significantly ($F(2,34) = 21.60$; $p < .001$) when followed by one, two and three constituents (verb included). The increase in duration between one and two arguments was larger than the one between two and three arguments. We have no explanation for the fact that the realization of the nominative takes more time when there was a greater number of following constituents. As a reviewer suggests, the greater number of prosodic phrases might exert an influence here, but in a way that we do not understand.

[7]A reviewer observes that this effect looks like declination in the sense of Pierrehumbert (1980), i.e. a phonetic lowering effect independent of phonological downstep. An alternative explanation could be that downstep takes place in a compressed pitch range.

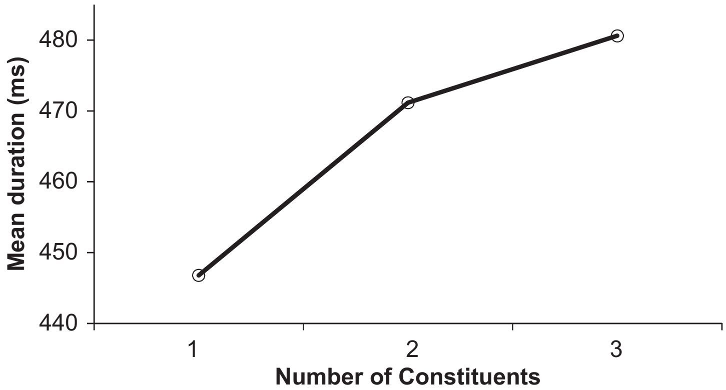

Fig. 9. Effect of number of constituents on the duration of the nominative; mean duration in ms is plotted against the number of following constituents (1, 2 and 3).

3.5. *Word order*

The last hypothesis related to information structure was on word order: a non-canonical word order (in which the nominative is not initial) may have an influence on the height of pitch accents (Hypothesis 5). This hypothesis was not confirmed: word order did not affect any of the values calculated in this study. As far as pitch is concerned, we compared the accents of the arguments in the unmarked patterns N<u>A</u>V and N<u>D</u>V with those in the marked patterns A<u>N</u>V and D<u>N</u>V. Sentences with the same arguments but in a different order did not differ in the peak f0 of their arguments. The f0 for A1 and A2 were not significantly different regardless of their nominative, accusative or dative case and regardless of their information status. In the longer conditions, again no difference could be found in the f0 of accented words. The values were similar in unmarked and in marked word orders.

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

3.6. *Constant final value and final f0 drop*

Hypothesis 6 predicted a constant final value for all our data, regardless of length and information structure. This hypothesis was confirmed by the data, and is thus in line with Liberman and Pierrehumbert (1984) and Grabe (1998). Figs. 5a–c and 3d and e illustrate the constant final low point of the speakers in our data. Only the values on an unaccented verb can be used here. Only in these cases, the maximum f0 measured by the Praat script corresponds to the lowest value of the sentence.

A much larger downward trend than predicted by downstep between the last narrowly accented argument and the first unaccented constituent was observed. This considerable drop in f0 was on average 80 Hz. Interestingly, not only the final low of the sentence but the drop itself was constant: 80–85 Hz. Paired-sample *T*-tests on the comparison of the final drop across different conditions showed no significant difference between narrow focus in different positions. But the low f0 value that was reached depended on which of the first, second or third arguments was narrowly focused. At the end of the sentence, the constant final value of 190 Hz or less was reached in all cases, but sometimes it needed more than one decreasing step to be achieved (see the values in Fig. 5a and b).

We also find a constant final value in all-new sentences in the realizations with an unaccented verb, although there was a significant difference between the all-new condition and the narrow focus conditions. The final drop was significantly larger for narrow focus ($t(17) = 6.58$; $p<.001$ for all-new vs. initial narrow focus; $t(17) = 7.79$; $p<.001$ for all-new vs. second argument focus; $t(17) = 7.60$; $p<.001$ 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 which involved final f0 drop.[8]

[8]We only found the final drop for sentences containing an unaccented verb, thus 101 instances out of a total of 348 all-new sentences.

Table 7
Pitch peak values on the first (A1), second (A2), and third (A3) arguments before focused constituents in all-new sentences (left column) and before non-focused, i.e. given, constituents (right column); in narrow focused sentences, underlined constituents are focused, the others given

	Before [+foc]		Before [–foc]
A1			
NV	264	NV	277
NDV	271	NDV	289
NAV	269	NAV	285
NDAV	284	NDAV	289
A2			
NDV	257	NDV	270
	257	NDV	275
NAV	255	NAV	269
	255	NAV	271
NDAV	253	NDAV	276
A3			
NDAV	260	NDAV	263
	260	NDAV	269

3.7. Further tonal effects

Following Hypothesis 7, all pitch values of high tones in our data could be interpreted as a combined effect of syntactic structure and focus/givenness partition. Our data cannot confirm this hypothesis. Some values that were calculated on the tones themselves, as for example the raised values on the upstepped argument in the all-new patterns, or the different values of a given argument which vary as a function of the adjacency or non-adjacency of the narrowly focused constituent, could not be explained by syntax and information structure alone. In other words, the f0 on some pitch-accented words could be explained neither in terms of regular downstep, nor by boosting due to narrow focus, nor by lowering due to givenness.

The first such value concerns a narrowly focused pitch accent just before the final drop. In this environment, the pitch was raised, an effect which is not limited to the preverbal accent, but also affects an accent located immediately before an unaccented constituent, regardless of both sentence position and information structure. Table 7 makes some relevant comparisons from which it appears that the f0 of an accented argument is lower when it precedes a further accented argument (second column) than when it precedes an unaccented one (fourth column). We interpret the higher pitch peaks on the latter case as a result of H-tone raising. On average, the difference between a raised tone and a non-raised one was 17 Hz and was significant ($t(17) = -4.12$; $p < .01$ for the first argument; $t(17) = 5.03$; $p < .001$ for the second argument).

N of NDAV, i.e. a nominative followed by two more arguments, did not show a significantly raised f0 when the following dative was unaccented. In this case, there was only 5 Hz difference between the value of N in an all-new context and in a narrow focus.[9] The second place where the difference was only weak is A3 in NDAV.

Some of the raised values in Table 7 are indistinguishable from the effect of narrow focus. But two facts cast doubt on the view that raising of a high tone can always be explained as a result of narrow focus. First, as is visible from Table 7, A2 is raised both when it is narrowly focused (e.g. NAV), and when it is part of a pattern in which all arguments are focused with the exception of the verb (e.g. NAV). Second, H-raising also takes place in an all-new pattern, when the verb is unaccented and the preverbal argument is upstepped. The relevant values from Fig. 3a and d are reproduced in Table 8. The upstepped argument is on average about 20 Hz higher than in a downstep pattern. In this case, narrow focus plays no role at all.

The explanation advanced here is that of a purely tonal interaction between two adjacent tones. We think that the phenomenon of H-raising, which typically occurs when a high tone precedes a low tone in tone

[9]We suspect that a ceiling effect is at play in this case, but more investigations are needed to confirm this tentative explanation.

Table 8
Pitch peak values in Hz of a preverbal dative (D) or accusative (A) argument in all-new sentences where the argument occurs in an upstep pattern (left column) or in a downstep pattern (right column)

	Upstep pattern	Downstep pattern
NDV (D)	271	259
NAV (A)	277	251
NDAV (A)	280	258

Table 9
Pitch peak values in Hz of a prefocal argument (left column) compared to the same argument in the downstep pattern of all-new sentences (right column)

Narrow focus sentence		All-new sentence	
N**A**<u>V</u>	A = 239	N**A**V	A = 250
N**D**<u>V</u>	D = 238	N**D**V	D = 260
N**D**<u>A</u>V	D = 229	N**D**AV	D = 264
ND**A**<u>V</u>	A = 229	ND**A**V	A = 254

languages (see among others Connell & Ladd, 1990; Laniran & Clements, 2003; Xu, 1997 for H-raising in tone languages) plays a role in German as well.

The second values, which cannot be explained by syntax and information structure, can be understood as the reverse effect of H-raising. A high tone immediately preceding a raised high tone was lowered as compared with the same tone preceding a non-raised high tone. Several data sets illustrate this effect.

First, as was shown in Fig. 6 in Section 3.3.1, the lowest values for an initial (given) nominative were those in which it immediately preceded a narrow focus (column C). When the second argument was given as well (column D), the values of the nominative were higher by ca. 20 Hz on average across speakers. Paired-sample *T*-tests revealed that a comparison between the values of the two initial given nominatives produced a significant effect ($t(17) = -8.50$; $p < .001$). 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 9 compares the height of arguments immediately preceding a narrowly focused argument with similar ones in the downstep pattern of all-new sentences. The former values (left column) were surprisingly stable (on average between 229 and 239 Hz), and systematically lower than the values they had in an all-new sentence (right column).

Even if it is assumed that givenness is responsible for the low values in the left column of Table 9 such an interpretation is not available for the third set of data: an argument preceding an H-raised tone in an all-new sentence. The relevant values are shown in Table 10. Paired-sample *T*-tests show that, in an all-new sentence, a constituent preceding an accented constituent was significantly lower than a corresponding constituent in a regular downstep pattern ($t(17) = 7.06$; $p < .001$). The difference was again 17 Hz or more on average. In this case, as well, an argument preceding a raised value was lowered as compared with a downstep pattern. Since the all-new status of the constituents was identical in each case, it must be assumed that it was the tonal make-up of the phrase that was responsible for the difference.[10]

4. Discussion

The data on tone scaling presented in Section 3 confirm a view of the tonal structure of German as largely determined by syntax and information structure. But they also revealed that syntax and information structure

[10]The discussion of Table 7 is also relevant. It was shown there that downstep between A1 and A2 before a raised argument is larger than downstep before a narrowly focused verb. In the latter configuration, no lowering can take place, because the verb is never H-raised (see discussion).

Table 10
Comparison of pitch peak values in Hz of arguments occurring in upstepped and downstepped patterns; in the case of preverbal arguments, the argument preceding this argument has been measured; in the case of the verb (V), the preverbal argument has been measured

Condition	Upstep pattern on preverbal argument	V	Downstep pattern on preverbal argument	V
NAV	N = 260	A = 245	N = 276	A = 251
NDV	N = 264	D = 242	N = 283	D = 259
NDAV	D = 257	A = 230	D = 275	A = 258

are not sufficient to explain all patterns of variation, and we proposed to add a third component to pitch accent scaling: interactions between the tones themselves. In this section, it is shown how the three components entering tonal scaling are integrated into a model of German intonation.

The prosodic phrasing of the experimental sentences is illustrated in (9) and (10). Every argument projects its own p-phrase. An all-new sentence can integrate the verb in the preceding p-phrase, as in (9a), or the verb forms a separate p-phrase, as in (9b) (see Fuchs, 1976; Gussenhoven, 1992; Jacobs, 1993). In such a sentence, every p-phrase has a head, realized as a pitch accent.

(9) a. [Weil der HAMMEL]$_P$ [den HUMMER eingeladen hat]$_P$
because the sheep the lobster invited has
'Because the sheep invited the lobster.'
b. [Weil der HAMMEL]$_P$ [den HUMMER]$_P$ [EINGELADEN hat]$_P$

In (10), a sentence with narrow focus and postfocal deaccenting is shown. Subscript F stands for focus, and G for givenness is present. Two phrasing structures are possible. Phrasing (10a), consisting of only one prosodic phrase, respects the assumption that every prosodic phrase should have its own head. Since the object and the verb are deaccented, they do not form a prosodic phrase, and are integrated into the phrase of the accented subject. In (10b), the phrasing based on syntax is kept. Only the height of pitch accents is changed, which are compressed in the second phrase. We have no clear argument that would allow us to choose between the two structures. Further research is needed on the issue of the phrasing of postfocal material, and on the nature of the deaccenting.

(10) a. [(Weil der HAMMEL)$_F$ (den Rammler eingeladen hat)$_G$]$_P$
b. [(Weil der HAMMEL)$_F$]$_P$ [(den Rammler eingeladen hat)$_G$]$_P$
'Because the sheep invited the buck.'

The tonal pattern of the sentences as illustrated in (11) contains prenuclear rising accents, which are transcribed as bitonal L*H$_P$ sequences, and nuclear falling accents which are transcribed as H*L$_I$. L* is a low tone associated with an accented syllable, and H* is the high pendant. H$_P$ is the boundary tone of a prosodic phrase (p-phrase), and L$_I$ the boundary tone of an intonation phrase (i-phrase). Autosegmental–metrical transcription was originally proposed by Pierrehumbert (1980) for English (but see Féry, 1993; Grabe, 1998; Grice, Baumann, & Benzmüller, 2005; Truckenbrodt, 2002 among others, for adaptation to German).[11] A low boundary tone (L$_I$) is associated both with the syllable following H* and with the end of the sentence. The low stretch characterized by this tone extends backwards from the end of the sentence to the nuclear accent. This is in line with the OT implementation by Gussenhoven (2004) for English.

L* H$_P$ L* H$_P$ H* L$_I$
(11) [Weil der RAMMLER]$_P$ [dem REIHER]$_P$ [den HUMMER vorgestellt hat]$_P$
'Because the buck introduced the lobster to the heron.'

[11]Note that the transcription conventions differ considerably between researchers due to different systems of intonational phonology. See Kügler (2007, 25ff.) for a thorough discussion of different approaches to German intonation.

We consider regular downstep as the unmarked tonal realization (e.g. Féry, 1993; Truckenbrodt, 2002, 2004), in which unmarked prosody causes relatively small but regular descending steps, on average between 17 and 29 Hz. This pattern appeared in approximately half of the all-new sentences of our corpus ($n = 159$, 45.7%), as discussed in Section 2.1. Downstep also appeared in further environments: in a sequence of two prefocal pitch accents (see Section 3.3.1), and between A1 and A2 when all the arguments are focused with the exception of the verb (Section 3.2). In short, downstep features regularly in a sequence of two or more arguments when they have the same informational status, i.e. when they are equally focused or when they are equally given. But downstep can be smaller or larger depending on the tonal context in which it occurs. For given constituents, the downstep steps are smaller since the scaling of the pitch accents is already reduced due to givenness, while for focused constituents these steps can be larger due to the expanded pitch range.

We analyze the high part of a rising accent L^*H_P as the boundary tone of the p-phrase it ends; see (11) above for an illustration. The high parts of the rising contour are maximally as high as the reference line (top line) of their own domain. Downstep creates new (and downstepped) reference lines reached by the high tones. We assume that downstep takes place across p-phrases, as well as inside of p-phrases, in disagreement with Beckman and Pierrehumbert (1986), but in line with Bruce (1977), Clements (1990), Ladd (1990), van den Berg, Gussenhoven, and Rietveld (1992), Truckenbrodt (2002), Féry and Truckenbrodt (2005), among others (see Féry & Ishihara, to appear, for an elaborate model of German prosody along these lines). The register top and bottom lines are the intervals of the speaker's voice range, within the limits of which pitch excursions are scaled. Thus, when a high boundary tone is raised, it means that it is the register of the p-phrase delimited by this boundary tone that is raised. An alternative explanation for downstep has been advanced by Pierrehumbert (1980) and Liberman and Pierrehumbert (1984) for intonation languages, in which a medial L tone affects a following H in such a way that the affected H is lower than a preceding H. This effect takes the form of a left-to-right progressive assimilation. It is assimilatory because a low tone lowers a following high tone, thus rendering it more similar to itself. However, the assimilation-driven account of downstep cannot explain instances of downstep with no intervening low tones, as is the case in our data, as in (11) for example, where the high phrasal boundary tone H_P and the high tone of the pitch accent H* are in a downstep relationship to each other. Assuming downstep as a register effect related to prosodic domains includes cases such as (11). Further, downstep as a register effect taking place in different prosodic domains allows to account for embedded effects of downstep, as well as upstep.

The effect of information structure on the height of accents and boundary tones was discussed in Sections 3.2 and 3.3. The value of the high boundary tone of non-final p-phrases, as well as the value of the high pitch accent of the final p-phrase were measured and compared to each other. It was found that information structure affects the top lines of prosodic domains. Narrow focus raises it, and as a consequence the high part of L^*H_P is higher, and givenness lowers it prefocally and compresses it considerably postnuclarly. The high tone values were systematically higher in a constituent with a narrow focus than in its correspondent with an all-new context or in a given context. The effect of downstep was never cancelled, as a comparison between the values of a narrowly focused first, second and third arguments revealed. In short, the combined effects of syntax and information structure go a long way in the explanation of the high tone values of our data. However, other effects were acting on the scaling of tones, as well, to which we turn in the remainder of the discussion.

The final accent of the declarative sentences examined in the experiment was always a falling tone, which we represented with a high starred tone followed by the low boundary of the i-phrase, H^*L_I. The boundary tone L_I reached the bottom of the speaker's voice, a value we have called the 'constant final value.' In a sentence containing postfocal material, the final level was reached immediately after the last accent, or nearly so, otherwise it was attained on the final word of the sentence.

Final drop takes place in a considerable number of cases. This effect characterizes the steep fall between the last and low boundary tone L_I and the preceding upstepped high part of the pitch accent H*. It was regularly found in narrow focus, but also in 54% of the all-new sentences. We assume that the low tone L_I has a dissimilatory effect on a preceding high tone, the dissimilation being obligatory when the high tone stands for a narrow focus, and optional in an all-new sentence. The right-to-left dissimilation between the L_I and the H* causes an important difference in pitch. We found that the decrease in Hz taken by this final drop was remarkably stable. Since final drop happens independently of the information status of the constituents carrying these tones, it cannot be considered as an effect of focus.

Final drop should not be confused with Final Lowering, an effect which has been abundantly described in the literature on intonation languages (see Liberman & Pierrehumbert, 1984; Prieto, Shih, & Nibert, 1996; Truckenbrodt, 2004 for very interesting and elaborate analysis of Final Lowering in German).[12] Final lowering affects the last pitch accent in a series of accents, the final step is larger than expected by exponential decay of the decreasing values of the sequence of high tones. However, final drop is a different phenomenon, which relates to the register reference lines and not to accents. The steeper fall involved in final drop is not the last one in a series of accents, but it is the first in a series of unaccented words, or the only unaccented word, and the drop in f0 is much larger than the last step of a series of accents. The only places in our data where it would make sense to search for Final lowering are the sentences in which all constituents are accented, even the verb, since this is the only constellation in which downstep occurs with complete regularity. Yet, the effect of final lowering lies not in the scope of this paper, and thus we do comment on this issue.

Phonologically, we analyze the high value of the last H* before final drop as a dissimilation between a high tone and a following low tone. At the end of the intonation phrase, the final L_I exerts a local raising effect on an immediately preceding H*. The process of H-raising appears to take place in two steps. First, the final low tone, L_I, is scaled at the bottom line of the voice register, or close to the bottom, if it is not the last accentable constituent of the sentence. This L_I is also aligned with the syllable following the nuclear accent, creating in this way a flat and low postnuclear contour.[13] The second effect is the increase in pitch that L_I causes on the preceding H*, which then raises its value independently of the tones preceding it. It was shown in Section 3.7 that the value of the raised H* is dependent on final drop, as it is the difference between H* and L_I which matters.

The analysis of H-raising as a dissimilation process between two tones is supported by 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 and Clements (2003) find a dissimilatory H-raising effect in Yoruba: a low tonal onset raises the f0 of the preceding high tone (see also Connell & Ladd, 1990). Gandour, Potisuk, and Dechongkit (1994) also find an H-raising effect in Thai.

Alternative explanations propose that boosting of an accent is always an effect of narrow focus (Cooper et al., 1985; Eady & Cooper, 1986; Eady et al., 1986; Liberman & Pierrehumbert, 1984 for English; Kubozono, 1989 for Japanese). Clearly, our data refute this thesis since we find the same raising effect in all-new sentences. In the German sentences examined in this paper, H-raising 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.

In their study of tonal interactions in Yoruba, Laniran and Clements (2003) show that high tones are sometimes higher than predicted. 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 of an initial H in anticipation of a long high tone sequence. They express their proposal 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 tones will not penetrate the frequency band reserved for M tones' (2003, p. 232). In other words, H-raising guarantees that a downstepped H is not mistaken for a mid or even a low tone. However, their explanation for H-raising does not fit our data. Our speakers do not reset their voice in the same way as Yoruba speakers do. It is not clear whether there is a tonal domain reserved for low tones in German.

Upstep, as described by Truckenbrodt (2002), is not adequate for our data either. This phenomenon is limited to boundary tones of intonation phrases and it only affects medial i-phrases. The effect described here is a different one: it affects sentence-final bitonal accents, analyzed as pitch accents followed by boundary

[12]In a nutshell, Truckenbrodt proposes that final lowering is the effect of absence of raising of a high tone before a downstep tone. Since a sequence of high tones are in a downstep relationship, all are affected except for the last one.

[13]This may be a simplification, due to the observation that a non-final postnuclear word may be followed by decreasing steps, but this does not bear on the argumentation in this section.

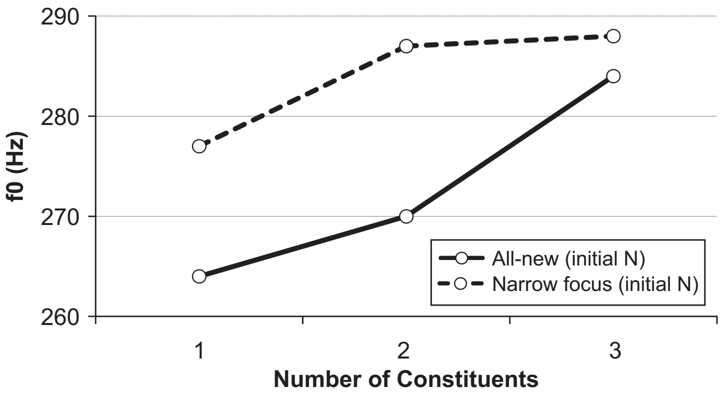

Fig. 10. Mean f0 of the initial nominative in an all-new sentence (solid line) and in a narrow focus sentence (dashed line) as a function of the number of following constituents.

tones of prosodic phrases. A last 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 reaches a point in the speaker's voice at which it can trigger the final drop. In our data, it is located at roughly 80 Hz above the point reached by the final drop. The raised H is, in other words, not scaled relative to the reference line of the first accent in the sentence. A more accurate approximation is that it is scaled to the last boundary tone, i.e. the constant final low. More research is needed to understand this phenomenon better, especially when final drop happens relatively early in the sentence.

The raised H* also has a dissimilatory effect from right-to-left on a preceding H, decreasing its value. The affected high tone is the second tone of a rising bitonal contour that is phonologically represented as L^*H_P. In other words, a starred H lowers the value of the preceding boundary tone of a p-phrase. This effect can be phonologically analyzed as a regressive dissimilation of an H* on a preceding H_P. The first high tone is rendered less similar to the second one. As in final drop and in H-raising, H-lowering takes place both when the raised H* is the narrow focus and in an all-new sentence. The sequence L^*H_P H^* L_i is the final sequence of tones in a declarative sentence.

This H-lowering can again be considered as a two-step process. First, the second high tone (the one acting on the preceding tone) is raised. Second, the first H tone (the one on which the action is exerted) is lowered and is thus dissimilated from its following tone.

Liberman and Pierrehumbert (1984) observed that a high tone preceding a boosted high tone is lowered, but, in contrast to our analysis, attribute this effect to information structure. We find, however, that the decisive factor is not so much the information structure, but rather the tonal structure of the sentence. H-lowering happens before a raised H independently of the status of the lowered tone as given or as part of an all-new sentence, and independently of the status of the raised tone as narrowly focused or as part of an all-new pattern.

Thus H*-raising increases at the same time the difference between this last H* and the preceding high tone. But, when the final verb is focused, no such effect occurs, and the difference between the last tone and the preceding one is not so large. A final narrowly focused verb has a lower pitch accent (ranging on average between 255 and 263 Hz) than a narrowly focused preverbal argument (268–271 Hz). The reason for the different behavior of the final verb is due to the fact that the verb, being the final constituent, is not followed by a deaccented constituent, and that a final drop is not needed in this context. There is no final drop, and consequently no H-raising. As a result, H-lowering does not apply either.

A final remark on the effect of different sentence lengths on the pitch scaling is in order. The average values of an initial nominative in a whole-focused sentence, and in a narrow focus are shown graphically in Fig. 10. They are clearly higher when the sentence contains more arguments.[14] In all-new sentences, there is a

[14]This effect can be compared with long-distance anticipation, which has been described for tone languages (Stewart, 1993 for Dschang and Ebrié, Rialland & Somé, 2000 for Dagara, to cite just a few) and to a lesser extent for intonation languages (see Thorsen, 1985 for Danish, and the rejection of preplanning by Liberman & Pierrehumbert, 1984; recently for Romance languages, Prieto, D'Imperio, Elordieta, Frota, & Vigário, 2006). In our data, we find evidence for soft preplanning in the sense of Liberman and Pierrehumbert (1984).

significant raising of the value of the initial nominative when the number of arguments increases. When the nominative is narrowly focused, this value rises between one and two arguments, but it does not increase significantly between two and three arguments.

We tentatively explain the absence of raising in the latter case in the following way: speakers have a register top line which cannot be raised *ad libitum*. In our data, a longer sentence or a raised nominative already reaches the ceiling. As a consequence, it is not possible to raise this top line even more. A confirmation of this view was noticed in Section 3.3. Raised initial nominatives because of narrow focus are significantly higher than the initial nominative in the all-new pattern, with the exception of NDAV, where the difference is not significant. First a raised nominative was only 5 Hz higher than a non-raised one, and second, it was at the same height as a nominative in a shorter sentence. Also in this case, a ceiling effect may be assumed that blocks further tone raising. However, this has to be shown experimentally.

5. Conclusion

Laniran and Clements' (2003) study of tonal interactions in Yoruba has significantly influenced the interpretation of the results of the present paper. Laniran and 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 f0 contour is the result of a compromise between different tendencies in the tonal pattern. The authors give a compositional analysis of tone scaling in which the melodic contour of an utterance is the result of a number of interacting factors. In the same way, we propose that the tonal scaling of high tones in a declarative German sentence is the result of different influences coming from syntax, information structure and tonal dissimilations.

The present work has examined tonal issues in German with the help of a production experiment in which 18 speakers uttered a total of 2277 sentences of the same syntactic structure, but with a varying 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 as has been suggested before for other Germanic languages. However, this experiment also showed that not all the data can be explained by information structure. Therefore, we propose an additional explanation to information structure, namely, that tones act on each other resulting in dissimilatory tonal effects. As for the effects of information structure, 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 out 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 appear to be 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 a tonal effect taking place across p-phrases: an L anticipatorily assimilates a following H tone, and thus lowers its value, but we have preferred to analyze it as an effect of downstepping the prosodic reference lines relative to each other. Second, final f0 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 all-new sentences, when the last high tone is H-raised. Third, H-raising, which has been described among others by Connell and Ladd (1990), and Laniran and Clements (2003) for Yoruba, and Xu (1997) for Mandarin Chinese, plays a crucial role as a component of the tonal pattern of German. Optional H-raising is triggered by a low tone exerting a dissimilatory regressive effect on a preceding high tone. We argued that raising a last accent allows the phenomenon of final f0 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 prefinal 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 f0 drop cannot be understood as final lowering as it involves a preceding raise on the last accent.

A side issue of this paper concerns the question of preplanning (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 providing some evidence for a mild form of preplanning, as a tone has an influence on a preceding one. H-lowering was shown to even involve a kind of double preplanning. In any case, our data appear to support the idea of 'soft preplanning' as argued for by Liberman and Pierrehumbert (1984).

Our investigation has necessarily been limited in scope 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.

Acknowledgments

This paper is part of the project A1 of the SFB 632 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, Jonathan Harrington, Robin Hörnig, Shin Ishihara, Ineke Mennen, Anja Gollrad, Gerrit Kentner, Lisa Selkirk, Hubert Truckenbrodt, Thomas Weskott, Wang Bei, Sabine Zerbian and two anonymous reviewers for helpful discussions and comments. The usual disclaimers apply. Our deep gratitude goes to Anja Arnhold, Kirsten Brock, Susanne Genzel and Esther Sommerfeld for technical support.

Appendix A

An overview of the speaker variation is shown in Table A1.

Table A1
Distribution of the realization of all-new sentence patterns by speakers; each of the four conditions were repeated five times ($5 \times 4 = 20$). Less than 20 implies that some f0 values could not be measured due to creaky voice

Speaker	Accented verb			Unaccented verb		
	Downstep	Raising on argument	Raising on verb	Downstep	Raising on argument	Total
1	15	2	1	1	1	20
2	1	1	6	1	9	18
3	2	2	8	4	3	19
4	3	1	7		5	16
5	6		9	1	4	20
6	6	3	11			20
7	4		7	3	1	16
8	2	1	13	1	3	20
9	9	1	2	5	3	20
10	15			3	2	20
11	8		8	3	1	20
12	11		8		1	20
13	12	1	3	3	1	20
14	14	1	4	1		20
15	2	1	2	7	6	18
16			14	3	3	20
17	4		2	3	11	20
18	4		8	1	6	19

References

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 P. Bosch, R. van der Sandt (Eds.), *Focus and natural language processing, Vol. 1, Intonation and syntax* (pp. 11–28). IBM Working Papers of the Institute for Logic and Linguistics.

Baumann, S. (2006). *The intonation of givenness*. Tübingen: Niemeyer.

Baumann, S., Grice, M., & Steindamm, S. (2006). Prosodic marking of focus domains—Categorical or gradient? In *Proceedings of speech prosody 2006* (pp. 301–304). Dresden: TVD Press.

Beaver, D., Clark, B. Z., Flemming, E., Jaeger, T. F., & Wolters, M. (2007). When semantics meets phonetics: Acoustical studies of second-occurrence focus. *Language*, *83*, 245–276.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, *3*, 255–309.

van den Berg, R., Gussenhoven, C., & Rietveld, T. (1992). Downstep in Dutch: Implications for a model. In G. Docherty, & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 335–367). Cambridge: Cambridge University Press.

Bierwisch, M. (1966). Regeln für die Intonation deutscher Sätze. *Studia Grammatica*, *7*, 99–201.

Boersma, P., & Weenink, D. (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., & Kenworthy, J. (1980). *Questions of intonation*. London: Croom Helm.

Bruce, G. (1977). Swedish word accent in sentence perspective. *Travaux de l'Institut de Linguistique de Lund* 12, Gleerup, Lund.

Büring, D. (2001). What do definites do that indefinites definitely 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.

Connell, B., & Ladd, D. R. (1990). Aspects of pitch realization in Yoruba. *Phonology*, *7*, 1–30.

Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastiveness in question–answer contexts. *Journal of the Acoustical Society of America*, *77*, 2142–2156.

Cruttenden, A. (2006). The de-accenting of given information: A cognitive universal? In G. Bernini, & M. L. Schwartz (Eds.), *Pragmatic organization of discourse in the languages of Europe* (pp. 311–356). Berlin, New York: de Gruyter.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America*, *80*, 402–415.

Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R., & Lotts, D. W. (1986). Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments. *Language and Speech*, *29*, 233–251.

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

Féry, C. (1993). *German intonational patterns*. Tübingen: Niemeyer.

Féry, C., Fanselow, G., & Krifka, M. (Eds.). (2006). *The notions of information structure. ISIS 6*. University of Potsdam.

Féry, C., & Ishihara, S. (2005). *Interpreting second occurrence focus*. Ms. Potsdam.

Féry, C., & Ishihara, S. (to appear). How focus and givenness shapes prosody. In M. Zimmermann, C. Féry (Eds.), *Information structure from different perspectives*. OUP.

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

Féry, C., & Truckenbrodt, H. (2005). Sisterhood and tonal scaling. *Studia Linguistica*, *59*, 223–243 (Special issue ''Boundaries in Intonational Phonology'').

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.

Gilles, P. (2005). *Regionale Prosodie im Deutschen. Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin, New York: de Gruyter.

Grabe, E. (1998). *Comparative intonational phonology: English and German*. Doctoral dissertation, Universiteit Nijmegen.

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

Grønnum, N. (1992). *The groundwork of Danish intonation*. Copenhagen: Museum Tusculanum Press.

Gussenhoven, C. (1992). Sentence accents and argument structure. In Iggy Roca (Ed.), *Thematic structure. Its role in grammar* (pp. 79–106). Dordrecht: Foris.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Jacobs, J. (1993). Integration. In M. Reis (Ed.), *Wortstellung und Informationsstruktur* (*Linguistische Arbeiten 306*). Tübingen, Niemeyer.

Jun, S. -A. (Ed.). (2005). *Prosodic typology. The phonology of intonation and phrasing*. Oxford: Oxford University Press.

Katz, J., & Selkirk, E. (2005). *Two types of focus, two types of prominence. Ms*. Amherst: University of Massachusetts.

Kubozono, H. (1989). Syntactic and rhythmic effects on downstep in Japanese. *Phonology*, *6*, 39–67.

Kügler, F. (2007). *The intonational phonology of Swabian and Upper Saxon (Linguistische Arbeiten 515)*. Tübingen: Niemeyer.

Kügler, F. (2008). The role of duration as a phonetic correlate of focus. In P. A. Barbosa, S. Madureira, & C. Reis (Eds.), *Proceedings of the Speech Prosody 2008 Conference* (pp. 591–594). Campinas, Brazil: Editora RG/CNPq.

Kügler, F., Féry, C., & van de Vijver, R. (2003). Pitch accent realization in German. In *Proceedings of the 15th ICPhS*, Barcelona, Spain (pp. 1261–1264).

Ladd, D. R. (1980). *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.

Ladd, D. R. (1983). Phonological features of intonational peaks. *Language*, *59*, 721–759.

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. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Laniran, Y. O., & Clements, G. N. (2003). Downstep and high raising: Interacting factors in Yoruba tone production. *Journal of Phonetics*, *31*, 203–250.

Lenerz, J. (1977). *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Narr.

Liberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff, & R. T. Oehrle (Eds.), *Language, sound, structure: Studies in phonology presented to Morris Halle by his teacher and students* (pp. 157–223). Cambridge, MA: MIT Press.

Maeda, S. (1976). *A characterization of American English intonation*. PhD thesis. Cambridge, MA: MIT Press.

Mayer, J. (1997). *Intonation und Bedeutung*. Doctoral dissertation, Arbeitspiere des IMS, Universität Stuttgart.

Menn, L., & Boyce, S. (1982). Fundamental frequency and discourse structure. *Language and Speech*, *25*, 341–383.

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

Pheby, J. (1980). Phonologie: Intonation. In K. E. Heidolph, W. Flämig, & W. Motsch, et al. (Eds.), *Grundzüge einer deutschen Grammatik* (pp. 839–897). Berlin: Akademie-Verlag.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Doctoral dissertation. Cambridge, MA: MIT Press.

Prieto, P., D'Imperio, M., Elordieta, G., Frota, S., & Vigário, M. (2006). Evidence for *soft* preplanning in tonal production: Initial scaling in Romance. In R. Hoffmann, & H. Mixdorff (Eds.), *Proceedings of speech prosody 2006* (pp. 803–806). Dresden: TUD Press.

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.

Rooth, M. (1985). *Associations with focus*. PhD dissertation. University of Massachusetts, Amherst.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, *1*, 75–116.

Schwarzschild, R. (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics*, *7*, 141–177.

Selkirk, E. (2006). The prosodic phrasing of FOCUS followed by focus: The case of English. In C. Gussenhoven, & T. Riad (Eds.), *Tones and tunes: Studies in word and sentence prosody*. Berlin: Mouton de Gruyter.

von Stechow, A., & Uhmann, S. (1984). On the focus-pitch accent relation. *Groninger Arbeiten zur Germanistischen Linguistik*, *25*, 228–263.

Stewart, J. M. (1993). Dschang and Ebrié 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. (1985). Intonation and text in standard Danish. *Journal of the Acoustical Society of America*, *77*, 1205–1216.

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

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

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*, 61–83.