Pitch accent realization in German

Frank Kügler, Caroline Féry & Ruben van de Vijver
University of Potsdam
E-mail: {kuegler, fery}@rz.uni-potsdam.de, ruben@ling.uni-potsdam.de

ABSTRACT

The realization of pitch accents in German, Dutch and English has been extensively studied. However, as far as we know, a systematic comparison is still lacking. The present paper is a partial presentation of results of a larger study to fill this gap. Reseachers in the Netherlands and in Great Britain are conducting the same experiments as the one presented here, and in the near future, we will be able to compare the three languages. This paper provides the results for German, giving first an account of the vertical scaling of prenuclear rising and nuclear falling accents, second, an account of the horizontal alignment of falling nuclear tones in narrow and broad focus, and, third, more detailed phonetic data on downstep.

1. INTRODUCTION

As far as the phonetic realization of pitch accents is concerned, several features, for example the horizontal alignment or vertical scaling of tones, were investigated in the past. Studies on the horizontal alignment of tonal targets provide evidence for a relatively stable anchoring of tonal targets on the segmental string of an utterance, thus supporting the configurational view of intonational phonology [1], e.g. [2] and [3]. For German, [4] has recently shown that the alignment of the F0 peak of a falling nuclear pitch accent depends on the position of the accented syllable in an intonation phrase: whereas it is aligned invariantly at the right edge of the syllable rhyme in non-final position, it is aligned at the left edge of the syllable rhyme in final position. Moreover, [5] has shown that the alignment also depends on the dialect: for sentences with broad focus he found that the F0 peak is aligned relatively early in Hamburg German, but relatively late in the accented syllable in Berlin German. A recent study has compared vertical scaling of tonal peaks in the urban varieties of Hamburg and Berlin [6]. The results were quantitatively grouped into normal, high, very high and downstepped peaks but no data on the scaling of accents were given.

In the framework of a larger research on the realization of pitch accents in Dutch, English and German, the German data have been recorded and analyzed. The main goal is to compare simple pitch accents in the nuclear and prenuclear position in short sentences. The material has been kept extremely simple, in order to facilitate future comparison between the languages mentioned. This study aims mainly at providing phonetic data on the vertical scaling of prenuclear and nuclear tones in German. Additionally, these data will also allow us to compare intial and final lows of intonation phrases in order to establish a mean starting point of speaking relative to the tonal space made by the pitch accents in an intonation phrase.

The second goal was to analyze the horizontal alignment of tonal targets. As has been mentioned, it is hypothesized that tonal targets are anchored on the segmental string in a relatively stable way. We will thus be able to account for early, medial or late peaks in order to compare these data with the realization of English and Dutch pitch accents in the future. Finally, a systematic comparison of high targets of prenuclear and nuclear tones will result in more detailed data on downstep.

2. METHODS

Recordings. Twelve native speakers of Standard German (variety spoken in the Berlin-Brandenburg region) were recorded on a portable DAT-recorder in a sound-proof booth. Method of data presentation was a PowerPoint file that the subjects could regulate at their own pace. The questions and the appropriate answers appeared on the screen and the speakers were asked to read aloud both the questions and the corresponding answers. The sentences were separated by at least two fillers. The test sentences were presented in a randomized order.

No. syllables	Proper names	Phonetic transcription	
1 syllable	Mo, Li	[ˈmoː]	[ˈliː]
2 syllables	Mona, Lina	[ˈmoː.na]	[ˈliː.na]
3 syllables	Ramona, Melina	[ra.ˈmoː.na]	[me.ˈliː.na]
verb	sah	[ˈzaː]	

Table 1: Speech materials used in the present study.

Speech Materials. Two sets of sentences were constructed: one set consisted of just one proper name, the second set of a subject, a transitive verb and an object. The subject was a proper name with one to three syllables, the verb was always *sah* 'saw', and the object was the same names as the ones appearing as subject (cf. Table 1). Examples are *Li sah Lina* 'Li saw Lina' and *Lina sah Ramona* 'Lina saw Ramona'. The complete set of names is given in Table 1. All sentences were elicited by questions. One type of questions like *Wen sah Mona?* 'Who did Mona see' forced narrow focus on the proper name in the one-word sentence. The other type of question *Was ist passiert?* 'What has happened?' forced broad focus on the double-name sentences. All names were combined with each other, resulting in 30 double-name sentences per speaker and a total of 360 sentences. The one-word sentences occurred only once resulting in a total of 12 x 6 (72) sentences.

Analyses. The materials were digitized at a sampling rate of 16 kHz, 16bit, and the signal analysis was executed with Praat (© 1992-2003 Boersma & Weenik). Speech materials were labelled semi-automatically at the level of syllable. Tonal labelling was conducted automatically, labelling the initial low and the following highest point in F0 of the first proper noun. In the one-word sentences the final low was labelled after the highest point of F0, in the double-name sentences a second high peak was labelled automatically in the domain of the verb plus object resulting in pitch peaks either on the verb or on the object (see section 3 for discussion about contour type). Subsequently, final lows on the object were labelled. As a result, three measurements per one-word sentence, and four measurements per double-name sentence were obtained.

3. RESULTS

Unsurprisingly, the tonal pattern of our sentences was a falling tone on the nuclear accent and a rising tonal pattern on the prenuclear one. We found two different contour types for the double-accent sentences: a hat pattern as shown in Fig. 1, and a two-accent peak pattern as shown in Fig. 2, see [7]. Measurements of the second F0 peak revealed that the location of the peak depends on the contour type: while in the hat pattern, the second peak is located before the nuclear word, i.e. on the verb *sah* ‘saw’, in the two-accent peak pattern the second F0 peak is located on the nuclear word, more precisely on the accented syllable (see section 3.2).

Additionally, an example of a one-word phrase is given in Fig. 3. The trisyllabic proper name was realized with an initial low followed by a rise up to the F0 peak which forms the high tonal target of the falling nuclear pitch accent of that phrase.

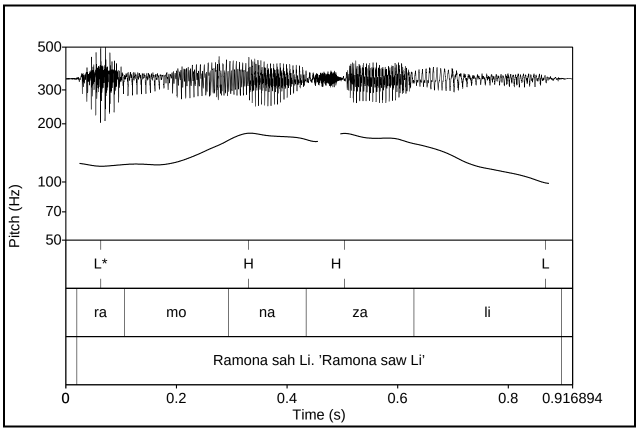

Figure 1: Example of the contour type “hat pattern” on a two-accent phrase as answer to the question *Was ist passiert?* ‘What happened?’.

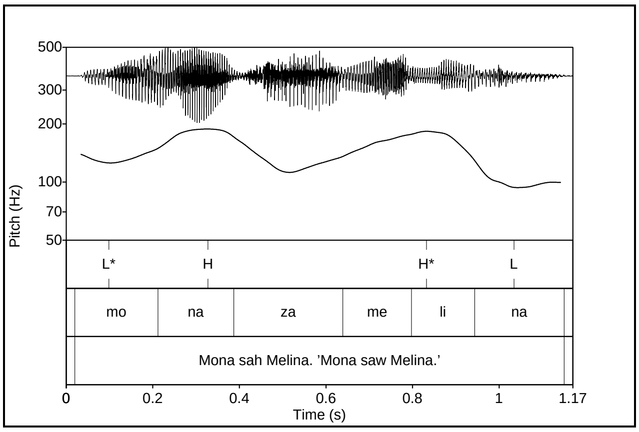

Figure 2: Example of the contour type “two-accent peak” on a two-accent phrase as an answer to the question *Was ist passiert?* ‘What happened?’.

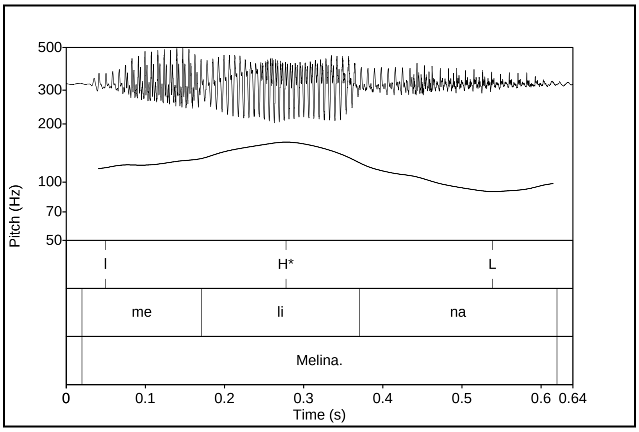

Figure 3: Example of a single accent phrase as an answer to the question *Wen sah Mona?* ‘Who saw Mona?’.

3.1 Vertical scaling of prenuclear and nuclear tones

The main goal of this research is to provide data on the vertical scaling of pitch accents in order to compare them with English and Dutch. Our ultimate goal is to explain differences in the amount of pitch change in the realization of pitch accents in these languages. This paper presents the results of the German data. Comparing the initial low of F0 of the one-word phrase (Tab. 2) with the low tonal target of the rising pitch accent of the two-accent phrase (Tab. 3), analysis of variance with the factor sentence type revealed a significant difference between the two ($F[1,403] = 85.43$, $p < 0.0001$). Post hoc t-tests showed significant differences in the initial low for each syllable type of the two sentence types ($p < 0.01$ for one syllable, $p < 0.05$ for two syllables, and $p < 0.005$ for three syllables).

27 tokens out of a total of 72 occurrences of the one-word phrases were not analyzed since, in these cases, no initial rise up to the high tonal target occurred. According to [4] this accent realization serves as a phonetic alternative to an initial rise, and thus cannot be analyzed as a different phonological category. Data of the remaining 45 tokens compared to the low target of L*H in the double-word sentence show an onset approximately 11 Hz higher than in

the one-noun sentences. Since no significant difference between the first F0 peak of each sentence type ($F[1,401] = 0.12$, $p > 0.05$) is observed, the difference in onset height (initial low) between the two sentence conditions is responsible for the significant difference in vertical excursion between the two conditions ($F[1,401] = 43.69$, $p < 0.0001$).

	Initial low			High of H*L			Vertical excursion		
	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n
1syll	137	31	14	155	32	14	18	14	14
2syll	138	38	13	157	37	13	19	8	13
3syll	119	19	18	156	38	18	37	28	18
total	130	36	45	156	36	45	26	22	45

Table 2: Vertical excursion of the rise to the nuclear high target in the one-word phrase (narrow focus).

	Low of L*H			High of L*H			Vertical excursion		
	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n
1syll	112	11	120	159	33	119	48	26	119
2syll	110	11	120	156	22	119	46	16	119
3syll	105	10	120	157	26	120	52	22	120
total	109	11	360	158	28	358	49	22	358

Table 3: Vertical excursion of the prenuclear rising pitch accent in the double-accent phrase (broad focus).

	High of H*L			Low of H*L			Vertical excursion		
	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n
1syll	150	29	23	97	19	23	53	28	23
2syll	154	34	24	92	12	24	62	34	24
3syll	157	39	24	94	24	24	62	35	24
total	153	34	71	94	19	71	59	33	71

Table 4: Vertical excursion of the nuclear falling pitch accent in the one-word phrase (narrow focus).

	High of H*L			Low of H*L			Vertical excursion		
	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n
1syll	141	20	120	95	12	120	47	16	120
2syll	147	21	117	94	13	117	53	18	117
3syll	143	25	120	93	12	120	50	23	120
total	144	22	357	94	12	357	50	19	357

Table 5: Vertical excursion of the nuclear falling pitch accent in the two accent phrase (broad focus).

Comparing the data of the nuclear falling pitch accent for each sentence type across syllable structures, no significant effect of the number of syllables can be found for the pitch peak ($F[2,68] = 0.21$, $p > 0.05$ for narrow focus, and $F[2,354] = 1.85$, $p > 0.05$ for broad focus). Thus, number of syllables has no effect on the height of the F0 peak. However, comparing the means of the F0 peaks across sentence type we find a significant effect ($F[1,427] = 9.53$, $p < 0.005$). The mean of the F0 peak of the falling accent in broad focus is about 9 Hz lower than in narrow focus condition. Since there is no significant effect of sentence condition on the low tone ($F[1,428] = 0.13$, $p > 0.05$) the difference in vertical excursion between the two sentence conditions ($F[1,425] = 10.16$, $p < 0.005$) is due to the difference of the high tone. On average, pitch excursion of the falling nuclear pitch accent is 14 Hz smaller in broad focus condition (19 Hz) than in narrow focus condition (33 Hz) (cf. Tables 4 and 5). Interestingly, the final low of each sentence is similar for both focus conditions on average (94 Hz, cf. Tables 4 and 5) with no significant effect, as mentioned before.

While the initial low depends on focus structure, as mentioned before, the final low does not. On average, a phrase in narrow focus starts 36 Hz higher than the final low, and in broad focus it only starts 15 Hz higher than the final low (cf. Tables 2 – 5).

3.2 Horizontal alignment of tonal targets

Here, we consider the tonal alignment of peaks of falling nuclear accents in narrow and in broad focus. The high part of the pitch accent was expected to be realized on the accented syllable. However, we found that the position of the pitch peak depends on contour type. Recall the patterns shown in Figs. 1 and 2, a hat pattern and a two-peak pattern. There seems to be free variation between the two since every speaker produces both patterns, though in different proportions. Concerning the pitch peak position, however, a unified picture across speakers arises: in the hat pattern the peak was consistently realized on the verb *sah* 'saw', while in the two-peak pattern it was consistently realized on the accented word. Table 6 gives the number of peaks realized on and before the nuclear word as well as the distance of the peak realization measured in seconds from the beginning of the nuclear word (minus means that the peak was realized before the nuclear word).

	In the nuclear word			Before the nuclear word			
	x̄ [sec]	SD	n	x̄ [sec]	SD	n	total
1syll	0.09	0.05	65	-0.12	0.06	55	120
2syll	0.12	0.07	79	-0.10	0.05	41	120
3syll	0.20	0.09	107	-0.09	0.05	13	120
total	0.15	0.09	251	-0.11	0.05	109	360

Table 6: Distance of the nuclear peak in relation to the beginning of the nuclear word, given in seconds.

Analysis of variance revealed a significant effect of the number of syllables with their peaks on the nuclear word ($F[2,248] = 47.50$, $p < 0.0001$), but no significant effect with peaks before the nuclear word ($F[2,106] = 2.32$, $p > 0.05$). This means that F0 peaks realized before the nuclear word are anchored at an equal interval independently of the number of syllables of the nuclear word. Peaks realized on the nuclear word, however, depend on the number of syllables: post hoc t-tests showed a significant difference between the distance from the begining of the nuclear word up to the pitch peak for trisyllabic words in comparison to di- and monosyllabic words ($p < 0.0001$ for both cases). This is not surprising since the trisyllabic names used in this study have word stress on the second syllable while mono- and disyllabic names have word stress on the first syllable. Thus we may conclude, that if the peak is realized in the nuclear word, it is, realized on its accented syllable. This is shown in Table 7 where it can be seen that the peak was realized on the pre-stressed syllable in trisyllabic words in 30 out of 107 cases (28%) in the double-name sentences and in 8 out of 24 cases (33.3%) in the one-word sentences, and on the post-stressed syllable in di- and trisyllabic words in 12 (15.2%) and 11 (10.3%) out of 79 and 107 instances in the double-name sentences, and in 2 (8.3%) and 4 (16.7%) out of 24 instances each in the one-word sentences.

Further, Table 6 shows that the majority of tokens (251 out of a total of 360) have been realized with the peak on the nuclear word. In particular, this is true for the trisyllabic names where only 13 out of 120 tokens were realized with the peak before the nuclear word.

stressed syllable	before n [%]	in n [%]	after n [%]	total [%]
1syll	-	100 / 100	-	100
2syll	-	84.8 / 91.7	15.2 / 8.3	100
3syll	28 / 33.3	61.7 / 50	10.3 / 16.7	100

Table 7: Number of tokens in per cent realized on the nuclear word divided into peaks realized before, in, and after the stressed syllable. The first number corresponds to the double-name sentences, the second number to the one-word sentences.

3.3 Downstep

Analysis of variance with peak as a factor revealed no significant effect of the number of syllables ($F[2,246] = 0.39, p > 0.05$ for H1 in the hat pattern, $F[2,246] = 0.61, p > 0.05$ for H2 in the hat pattern, $F[2,104] = 1.04, p > 0.05$ for H1 in the two-peak pattern, $F[2,104] = 0.77, p > 0.05$ for H2 in the two-peak pattern). Comparison between H1 and H2 per contour type, however, showed a significant effect ($F[1,212] = 13.20$, $p < 0.0005$, $F[1,496] = 43.07$, $p < 0.00001$). Post hoc t-tests showed the same significant effect for all number of syllables ($p < 0.05$ for all structures in both contour types) except for the hat pattern with only one syllable in between ($p = 0.061$). Generally, we can observe that the high tones were downstepped relatively to each other in all cases except in sentences with only one syllable between the pitch accents in the hat pattern, and in all cases in the two-peak pattern. In the hat pattern, the second peak is downstepped twelve Hz on average, in the two-peak pattern, it is downstepped 14 Hz on average (cf. Table 8).

No syllables between H's	H1 (prenuclear)			H2 (nuclear)		
hat pattern	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n
1	156	32	30	144	23	30
2	151	25	69	140	20	69
3	164	26	8	134	30	8
total	153	27	107	141	22	107
two-peak pattern	x̄ [Hz]	SD	n	x̄ [Hz]	SD	n
1	163	35	40	148	20	40
2	159	26	145	144	22	145
3	159	27	64	144	26	64
total	159	28	249	145	23	249

Table 8: Means of the two peaks in the double-name sentences given in Hz per contour type.

4. CONCLUSIONS

The present study provides three sets of results. First, phonetic data on the vertical scaling of pitch accents in German. We found a coherent tonal level for the initial peak of an intonation phrase both in narrow and broad focus conditions. Second, the pitch level of the second peak seems to be invariantly downstepped (about 12 to 14 Hz) independently of contour type. In a narrow focus condition, speakers tend to start higher than in a broad focus condition, resulting in a smaller rise up to the accentual peak. As far as nuclear falling accents are concerned, the fall from the peak to the phrase final low is smaller in the broad focus condition than in the narrow one. This could be attributed to a difference in accent type since, in broad focus, the second accent, i.e. the falling accent, is downstepped relatively to the first accent. The final low, however, is realized at about 94 Hz in both sentence types. As for the horizontal alignment of pitch peaks, our third set of results, we found that accentual peaks were realized on the accented syllable in the majority of cases. The results presented here on German will contribute to a comparison between English, Dutch and German.

ACKNOWLEDGMENTS

We would like to thank Laura Herbst for assistance.

REFERENCES

[1] R. Ladd, *Intonational Phonology*, Camb.: CUP, 1996.

[2] R. Ladd, I. Mennen, and A. Schepman, "Phonological conditioning of peak alignment in rising pitch accents in Dutch", *JASA* 107, 2685-2696, 2000.

[3] F. Kügler (submitted) "The Phonology and Phonetics of Rising Pitch Accents in Swabian", in *Regional Variation in Intonation*, P. Gilles and J. Peters, Eds. Tübingen: Niemeyer, *to appear*.

[4] E. Grabe, *Comparative Intonational Phonology: English and German*. Nijmegen, PhD-thesis, 1998.

[5] J. Peters, "Fokusintonation in regionalen Varietäten des Deutschen.", Ms. Potsdam University.

[6] P. Gilles, "Die Intonation final fallender Nuklei. Eine kontrastive Untersuchung zum Hamburgischen und Berlinischen," *Germ. Ling.* 157-158, 167-200, 2001.

[7] C. Féry, *German intonational patterns*, Niemeyer, Tübingen, (= Linguistische Arbeiten 285), 1993.