Resistance is futile – The intonation between continuation rise and calling contour in German

Oliver Niebuhr

Department of General Linguistics, ISFAS, Christian-Albrecht-University, Kiel, Germany
niebuhr@isfas.uni-kiel.de

Abstract

German knows two plateau-based phrase-final intonation contours: the high level plateau of the continuation rise and the descending plateau sequence of the calling contour. They occur within a narrow scaling range of only a few semitones. The paper presents production and perception evidence for a third plateau-based phrase-final intonation contour inside this narrow scaling range. The new plateau contour shows a F0 decrease of between 1-3 st (in the form of a slightly declining plateau or a descending plateau sequence), involves additional lengthening of the vowels underneath the plateau, and occurs when resistance is futile, i.e. when speakers signal that they finally, but reluctantly, give in to a demand of the dialogue partner. Phonological implications are briefly outlined.

Index Terms: intonation, plateau, rise, German, stylization.

1. Introduction

The end of a phrase is intonationally particularly rich. The Kiel Intonation Model (KIM) for German [1] and its labelling system ProLab [2], for instance, distinguish five different phrase-final intonation movements: a terminal fall, a non-terminal fall, a high level plateau, a small and shallow rise, and a large and steep rise until the upper limit of the speaker's modal voice range. Additionally, there is the stylized intonation [3], i.e. a stepped drop in pitch from a high to a mid-high plateau, which is also referred to as chanted call contour, vocative chant, or calling contour [4,5] (the term calling contour will be used henceforth). If we then also include the concave vs. convex contrast in phrase-final rises [6], complex contours like fall-rise sequences, and so-called pseudo-terminal contours [7], we quickly end up with more than a dozen formally and functionally different phrase-final intonations in German.

The autosegmental-metrical (AM) phonology is not as diverse as that of the KIM. The most prominent AM model for German, GToBI [8], only distinguishes between five different phrase-final intonations that are represented by the boundary tones L-%, H-%, H-^H%, L-H%, and !H-%. The latter boundary tone denotes the calling contour.

Irrespective of the differently sized inventories of phrase-final intonation categories, KIM and GToBI agree in postulating two different phrase-final intonations that are built from plateaux. They will briefly be characterised in the following.

First, there is the high level plateau or H-%. It is preceded by a rising pitch accent. The sequence of rise and high plateau is frequently referred to as continuation rise (Fig.4). In fact, it was found by [9] in a label-based analysis of the Kiel Corpus of Spontaneous Speech that 92-94% of all naturally produced continuation rises occurred at turn-internal phrase boundaries. So, the term 'continuation rise' seems to be fairly adequate, even though it is presumably not the ultimate characterisation of the communicative function of this first plateau contour.

The second type of plateau-based phrase-final intonation is the calling contour or !H-%. The formal and functional aspects of the calling contour have been subject of intense debate for a long time. However, hardly any experimental study has been conducted so far. The tenor of the mostly descriptive analyses may be summarized as follows: Stylization is not an isolated phenomenon, but a process that can be superimposed on different intonation patterns and hence generates many subtypes of stylized contours (including stylized rises) [3,10,11]. Yet, the calling contour in the form of "a 'stepping down' from one […] level pitch to another" [3] is clearly the prototypical type of stylization (Fig.4). How the step down in between the two plateaux of the calling contour is aligned with the segmental string is influenced by the underlying syllable structure [5]. The scaling of the step can, but need not be the often assumed interval of a minor third, i.e. 3 semitones. Rather, the scaling can vary substantially as a function of speaker and distance to the dialogue partner [12]. However, it mainly ranges between 2-6 semitones. Each plateau may start with a lengthening of the initial syllable [13]. With regard to communicative function, many studies agree that calling contours or stylizations in general are attention-seeking or convey a kind of contact signals. That is, they are used to establish or consolidate the communication channel between speaker and hearer, which, however, need not involve a large physical distance between them. Independent of the physical distance, calling contours bridge a personal distance between speaker and hearer, which is why they frequently occur in connection with routine matters and shared conventions between speaker and hearer [3,11,12,14,15].

The aim of the present paper is to provide initial production and perception evidence for a third plateau-based phrase-final intonation that is phonetically intermediate between the continuation rise and the calling contour. Although the study was conducted for German, it is very likely that the same phrase-final intonation occurs with the same meaning also in many other languages, at least in Western Germanic languages like English and Dutch.

2. Production data

The investigated phrase-final intonation contour was first discovered in 2009 during one of my prosody lectures at the University of Kiel. The students trained intonational labelling on the basis of audio recordings from enacted court cases in German TV, as a student objected that the calling contour label would be inadequate for a given phrase-final intonation. In fact, the problematic case showed only a very short initial plateau and only a marginal drop in F0 before the final plateau. Also the outlined function of the calling contour did not agree with the semantic-pragmatic context of the utterance. In addition, since the problematic case occurred turn-finally and involved a visible and perceptible decrease in pitch, labelling a continuation rise would have too been inadequate.

Inspired by this single-case study, we searched the entire TV recordings for formally and functionally similar phrase-final intonations and came up with 18 additional examples. PRAAT analyses (www.praat.org) of two of these examples are shown in Figures 1(a)-(b). All 19 examples were prosodically analysed in PRAAT with respect to their F0 and duration characteristics. The semantic-pragmatic contexts in which the examples occurred were also comparatively analysed.
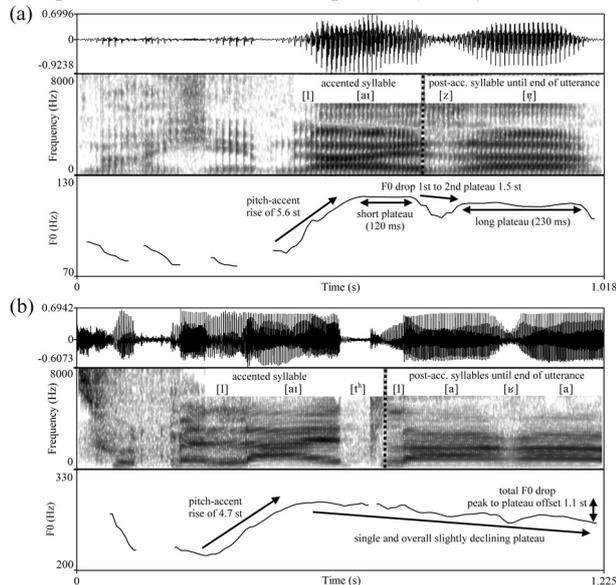

Figure 1: *Representative examples of the analysed phrase-final contours, realized after a rising pitch accent (a) as a sequence of short and long plateau in the utterance "Dann mach' ich eben leiser" (Well, all right, I turn it down), and (b) as a single slightly declining plateau in the utterance "Es tut mir leid, Lara" (I am sorry, Lara).*

The intonational characteristics of the 19 examples may be summarized as follows. The phrase-final contour is preceded by a pitch-accent F0 rise that starts close to the onset of the accented syllable (on average 28 ms after syllable onset) and peaks towards the end of the accented vowel (on average 33 ms before vowel offset). After the pitch-accent rise, there is in most cases (cf. Fig.1a) a short plateau until the end of the accented syllable. This first and short plateau is followed by a second and long plateau. It starts immediately after a sharp but tiny F0 drop at the onset of voicing in the first post-accented syllable and continues across all following syllables (2-5 in the analysed examples) until the end of the phrase. In some cases (cf. Fig.1b), the phrase-final intonation contour is realized as a single and overall slightly declining plateau that is directly attached to the pitch-accent rise.

Irrespective of whether the phrase-final contour consists of a single declining plateau or of a descending sequence of short and long plateau, F0 only decreases between 0.5 and 2.5 semitones (on average 1.47 semitones). The F0 decrease was measured in terms of the distance between the pitch-accent peak and – as in Fig.1b – the end of the single plateau, or – as in Fig.1a – the initial F0 value of the second plateau. Although the F0 decrease varies in a narrow range of just about 2 semitones, it is even possible to explain a part of this variation by the range of the preceding pitch-accent rise. That is, the larger the pitch-accent rise, the larger is also the F0 decrease in the phrase-final contour (r=0.61; df=17; p<0.01). As regards duration, all vowel nuclei in the sequence of pitch-accent rise and phrase-final contour are lengthened relative to reference

values from German spontaneous speech [16]. The lengthening is stronger for the unaccented than for the accented vowels and on the whole also impressionistically very salient. For example, the accented diphthongs of "leiser" and "leid" in Figures 1(a)-(b) are about 300 ms long. The following unaccented monophthong durations are between 300-450 ms.

The semantic-pragmatic contexts in which the 19 examples occur suggest that speakers use the phrase-final contours when *resistance is futile*, i.e. when they signal that they eventually, but reluctantly, give in to a demand of the dialogue partner. Accordingly, the utterance in Fig.1(a) was produced by a male teenager who reluctantly gave in to his parents and turned his music down. Example 1(b) comes from a convicted woman who reluctantly followed her lawyer's instruction and apologised to the complainant ("Lara") for having lied to her. The other 17 examples were produced under very similar semantic-pragmatic circumstances.

3. Perception experiment

It may be assumed that F0 changes – even decreases – of about 1 semitone and more are detectible in speech, particularly when they occur across long F0 stretches or between plateaux [17]. Therefore it was not the primary aim of the perception experiment to examine whether the contour found in the production data would be discriminable from the continuation rise on the one side and the calling contour on the other. Rather, the questions were whether (1) a stimulus continuum in the narrow intonational space between continuation rise and calling contour would in fact yield an additional, previously undiscovered communicative function, and whether (2) this function would be adequately outlined by the speaker's reluctant consent to give in to a demand of the dialogue partner.

3.1. Method

The stimulus generation started from one of the naturally produced utterances of our TV recordings: "Es tut mir leid, Lara", cf. Fig.1(b). The utterance tune was stylized in PRAAT at its major turning points (using a stylization threshold of 2 semitones). Then, the phrase-final intonation of the single slightly declining plateau was replaced by a sequence of short and long plateau, cf. Fig.2. The short plateau spanned the glide of the diphthong [aɪ] in the accented syllable "leid" (sorry). The long plateau extended over the following 2 syllables of the proper name "Lara" until the end of the utterance. Since both the single slightly declining plateau and the sequence of short and long plateau were found in the production data and are also impressionistically hardly distinguishable, it was reasonable to assume that both variants are equivalent representatives of the investigated phrase-final contour so that the replacement does not undermine the validity of the results. The use of a plateau sequence facilitated creating a stimulus continuum between continuation rise and calling contour.
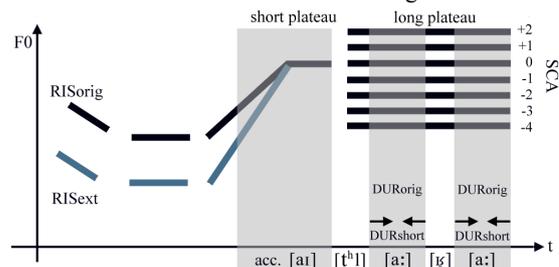

Figure 2: *Summary of the F0 and duration manipulations for the variables SCA, RIS and DUR.*

The PSOLA manipulations that were made for the base stimulus resulted from a cross-combination of three independent variables. The manipulations are summarized in Figure 2.

SCA: The scaling of the long plateau on "Lara" was varied in 7 steps of 1 semitone. Relative to the pitch-accent peak that transitioned into the first plateau on the accented syllable "leid", the long plateau had either same scaling, a 1-2 semitones higher scaling, or a 1-4 semitones lower scaling. So, the 7 steps may be called +2, +1, 0, -1, -2, -3, -4.

RIS: The range of the pitch-accent rise on "leid" was varied in two steps. One step corresponded to the original range of 4.7 semitones (RISorig). In the other step (RISext), the range was extended to 8.7 semitones by lowering the rise onset together with the utterance-initial F0 movement. The RIS variable was motivated by the correlation of pitch-accent rise and phrase-final contour scaling in the production data.

DUR: The durations of the two [a:] vowels within the long plateau on "Lara" were varied in two steps. One step corresponded to the original, lengthened durations (DURorig). In the other step (DURshort), both vowels were shortened by about 50% so that they their durations roughly corresponded to the normal average durations of unaccented [a:] vowels in non-final and final syllables of German spontaneous speech. The shortening was based on a triangular duration-point configuration in PRAAT so that the acoustic parameter transitions at the vowel boundaries remained widely unaffected.

The manipulations and PSOLA resyntheses yielded a total of 7x2x2= 28 stimuli. The perception experiment consisted of 7 copies of each stimulus. The resulting 196 stimuli were presented to listeners in an overall randomized order via loudspeakers in a sound-treated room at the Deptartment of General Linguistics in Kiel. Each stimulus was introduced by a short bleep and followed by a 4 seconds pause, during which the listeners judged the stimulus.

Listeners judged the stimuli with respect to their communicative functions in a 3 AFC task. Judgments were made by pressing buttons on a box in front of the listeners. The listeners were told that intonations change the communicative functions of utterances and that the present experiment would address this fact on the basis of the utterance "Es tut mir leid, Lara". Then, the listeners received written explanations and pictures that were to familiarize them with the communicative functions of the continuation rise, the reluctancy contour, and the calling contour (without using these or any other metalinguistic labels or phonetic descriptions of intonational categories). In brief, the illustrated explanations enabled the listeners to apply three judgment categories to the stimuli: routine excuse (= calling contour), reluctant excuse (= reluctancy contour), and introductory excuse that leads over to a justification (= continuation rise). Prior to the actual experiment, the listeners were familiarised with the judgment procedure and the stimulus spectrum by means a subset of 14 stimuli that were randomly selected for each listener.

The group of listeners consisted of 24 native speakers of German. They were between 21-27 years old and students of informatics or Germanistics. None of them suffered from hearing disorders or had taken part in a prosody lecture.

3.2. Results

Since all stimuli were identical except for the phrase-final F0 and duration manipulations, all significant changes in judgment behaviour must be due to these manipulations. On this basis, the results, which are summarized in Figure 3 in terms

of judgment percentages for each stimulus condition, show firstly that the SCA variable in the form of the 7-step scaling continuum yielded a clear threefold judgment change for each of the four cross-combined RIS and DUR levels. Secondly, it can also be seen that the RIS and DUR levels themselves caused different judgment patterns. Three-way repeated-measures ANOVAs were used to statistically test the effects of SCA, RIS, and DUR on the judgments. SCA was subdivided into three groups with reference to the naturally produced scaling ranges of the involved phrase-final intonation contours: +2 to 0 = group A; -1 and -2 = group B; -3 and -4 = group C. Judgment frequencies across all 24 listeners were the dependent variable. Separate ANOVAs were performed for each judgment category. Each ANOVA included multiple post-hoc comparisons (Sidak t tests) within the fixed factors. However, due to lack of space, only the final and most relevant results will be reported below.
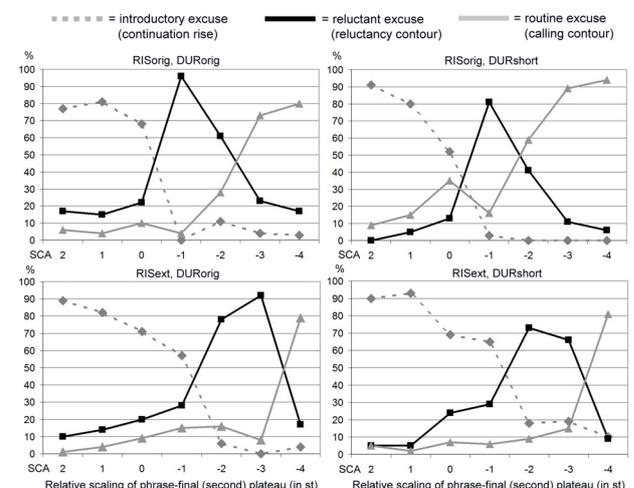

Figure 3: *Percentages of 'introductory excuse' (grey dots), 'reluctant excuse' (black line), and 'routine excuse' (grey line) judgments across the scaling continuum, displayed separately for each of the cross-combined RIS and DUR conditions. The data-point triplets at each scaling level represent 168 judgments.*

The ANOVA for 'reluctant excuse' yielded significant main effects of SCA ($F[2,46]=427.8$; $p<0.001$; $\eta_p^2=0.95$), RIS ($F[1,23]=17.07$; $p<0.001$; $\eta_p^2=0.43$) and DUR ($F[1,23]=81.04$; $p<0.001$; $\eta_p^2=0.78$). The three-way interaction was not significant. The effect of SCA is based on significant differences between all three groups A-C of the scaling continuum, reflecting the rising-falling identification curves in Figure 3. DURshort reduced the frequency of reluctant-excuse judgments compared with DURorig. The change from RISorig to RISext significantly increased the frequency of reluctant-excuse judgments. Figure 3 shows additionally that the RIS level influenced at which point the transition from introductory excuse to reluctant excuse took place in the scaling continuum. For RISorig, the transition to reluctant excuse occurred at -1, i.e. for the first descending F0 plateau, and it already changed again into a transition towards routine excuse when the phrase-final plateau was lowered by only one more semitone, i.e. to -2. In contrast, for RISext a plateau descent of -2 semitones was a prerequisite for the transition towards reluctant-excuse judgments, which then remained predominant until a plateau scaling of -3, i.e. a minor third. This fact is reflected in a strong interaction between RIS and SCA ($F[2,46]=$

148.95; p<0.001; η_p^2=0.87), whereas the interactions SCA x DUR and RIS x DUR were both not significant.

The results of the ANOVAs for 'introductory excuse' and 'routine excuse' were similar to those of 'reluctant excuse'. That is, there were significant main effects of SCA ('intro': F[2,46]= 4340.35; p<0.001; η_p^2= 0.99; 'routine': F[2,46]= 666.06; p<0.001; η_p^2=0.97) and RIS ('intro': F[1,23]=174.13; p<0.001; η_p^2=0.88; 'routine': F[2,46]=412.59; p<0.001; η_p^2= 0.95). The RIS x SCA interaction was only marginally significant for introductory excuses, but highly significant for the routine excuses (F[1,23]=174.13; p<0.001; η_p^2=0.88). Three-way interactions were not significant (in both ANOVAs).

Moreover, the factor DUR played a minor role. While it was completely irrelevant for the introductory-excuse judgments, shortening the vowels underneath the phrase-final plateau on "Lara" by 50% (from DURorig to DURshort) significantly *in*creased routine-excuse judgments (the same manipulation *de*creased reluctant-excuse judgments). However, this increase in routine-excuse judgments only occurred when the preceding pitch-accent rise had its original range of 4.7 semitones. Thus the RIS x DUR interaction was highly significant F[1,23]= 41.24; p<0.001; η_p^2= 0.64).

4. Discussion

If we assume that the 'introductory excuse' and 'routine excuse' judgments reflect the identification of continuation rises and calling contours, then the results of the perception experiment clearly support the production data from enacted court cases in suggesting the existence of a further form-function link within the plateau-based phrase-final intonation contours of German. In other words, this group of contours has three rather than just two members. The newly revealed contour is phonetically intermediate between the continuation rise and the calling contour (cf. Fig.4) and used when resistance is futile, i.e. when speakers signal that they eventually, but reluctantly, give in to a demand of the dialogue partner. The contour may thus be referred to as the *reluctancy contour* (Like continuation rise and calling contour, such functional labels are only working definitions that must be successively refined). One may assume that reluctancy contours are very often used by children, and my auditory observations in everyday life are in accord with this assumption. However, as our production study has shown, reluctancy contours also occur for adults and are anything but exotic for this speaker group.

As is illustrated in Figure 4, the plateau of the reluctancy contour differs from that of the continuation rise in that it does not remain at the level of the preceding pitch-accent peak. However, the F0 decrease is also smaller than for a calling contour; and while the latter must be realized as a stepping down from one level pitch to another, the same can, but need not apply to the reluctancy contour. Its plateau can also be slightly declining. That is, the F0 decrease relative to the preceding pitch-accent peak can either occur as a sharp F0 drop at the plateau-onset (as in the perception experiment) or as a successive F0 drop that is completed at the plateau-offset.

In addition, the reluctancy contour involves lengthening of the vowels underneath the F0 plateau. In line with [3] and unlike claimed by [13], no evidence was found that vowel lengthening was also relevant for calling-contour identification. In fact, the opposite was true. Vowel lengthening seems to reduce the identification of calling contours, cf. Figure 3.

Moreover, the perception evidence supports the production data in that the F0 decrease of the reluctancy contour

cannot be specified in absolute semitone values. It must be defined in relation to the range of the preceding pitch-accent rise. The larger the latter, the larger may also be the F0 decrease of the reluctancy contour.
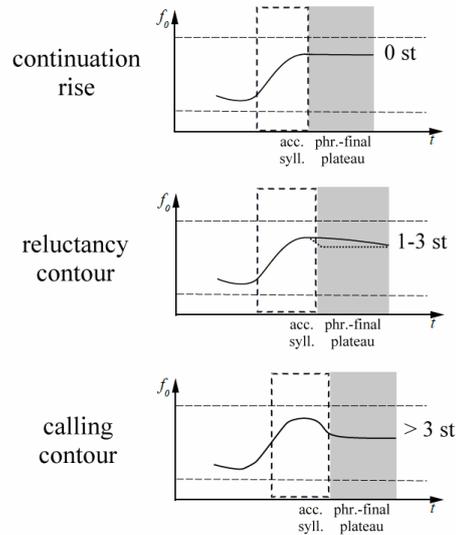

Figure 4: *The three plateau-based phrase-final intonation contours in German.*

It is reasonable to assume that a similar correlation between the scaling range of the final plateau and the preceding F0 interval also applies to the calling contour, since the scaling range of this contour is adjacent to that of the reluctancy contour and must be kept context-dependently distinct from it. The perceptual findings agree with this assumption, and hence pave the way for a follow-up study in the domain of speech production. Calling contours are typically associated with a minor-third interval, although the range of actually produced intervals [3,12,14] is much larger and varies considerably. It is likely that a pitch-accent related plateau scaling can help understanding, why the plateaux of calling contours are sometimes closer together and sometimes further apart along the frequency scale, in addition to the already known contextual and communicative factors.

Irrespective of this further question, it seems that the phrase-final intonation in German is still richer than claimed by major phonological models like KIM and GToBI. However, the assumed three-way contrast between continuation rise, reluctancy contour, and calling contour cannot be modeled in GToBI, since the !H-% representation is already reserved for the calling contour. One possible – and empirically underpinned – solution would be to outsource all (plateau) stylizations as a separate intonational mode that can then be applied to different kinds of intonation categories in order to change their communicative function in predictable but category-specific ways, cf. [10]. It is known for a long time that the timing of F0 movements relative to segmental landmarks can be very precisely controlled by speakers. The existence of three plateau-based phrase-final intonation contours within a narrow range of 3-4 semitones clearly shows that also the scaling dimension is controlled with remarkable precision.

5. Acknowledgements

6. References

[1] Kohler, K.J., "Paradigms in experimental prosodic analysis: From measurements to function", Sudhoff et al. [Eds], Methods in empirical prosody research, 123-152, Berlin: deGruyter, 2006.

[2] Peters, B., Kohler, K.J., "Trainingsmaterialien zur prosodischen Etikettierung mit dem Kieler Intonationsmodell KIM", URL: www.ipds.uni-kiel.de/kjk/pub_exx/bpkk2004_1/, 2004.

[3] Ladd, D.R., "Stylized intonation", Language 54, 517-40, 1978.

[4] Abe, I., "Call-contours", Proc. IVth ICPhS, 519-523, 1962.

[5] Schaeffler, F., Biersack, S., "Aspects of the timing of fundamental frequency in German chanted call contours", Phonum 9, 129-132, 2003.

[6] Dombrowski, E., Niebuhr, O., "Acoustic patterns and communicative functions of phrase-final rises: activating and restricting contours", Phonetica 62, 176-195, 2005.

[7] Ambrazaitis, G., "Between fall and fall-rise: substance-function relations in German phrase-final intonation contours. Phonetica 62, 196-214, 2005.

[8] Grice, M., Baumann, S., "Deutsche Intonation und GToBI", Linguistische Berichte 181, 1-33, 2002.

[9] Peters, B., Kohler, K.J., Wesener, T., "Melodische Satzakzentmuster in prosodischen Phrasen deutscher Spontansprache – statistische Verteilung und sprachliche Funktion", AIPUK 35a, 7-54, 2005.

[10] Dombrowski, E., "Semantic features of stylized vs. plain contours in German intonation", Phonetica, in press.

[11] Lindstrom, O., "Aspects of English intonation", Gothenburg: Acta Universitatis Gothenburgensis, 1978.

[12] Day-O'Connell , J., "'Minor third, who?': The intonation of the knock-knock joke", Proc. 5th Speech Prosody, 1-4, 2010.

[13] Gussenhoven, C., "The phonology of tone and intonation", Cambridge: CUP.

[14] Gibbon, D., "Perspectives of intonation analysis", Frankfurt: Peter Lang, 1976.

[15] Fagyal, Z., "Chanting intonation in French", U. Penn Working Papers in Linguistics 4, 77-90, 1997.

[16] Simpson, A.P., "Phonetische Datenbanken des Deutschen in der empirischen Sprachforschung und der phonologischen Theoriebildung", AIPUK 33, 1-233, 1998.

[17] Isacenko, A.V., Schädlich, H.J., "A model of standard German intonation", Mouton: The Hague, 1970.