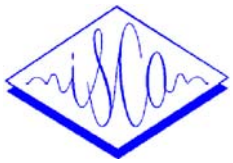


Speech Prosody 2002
Aix-en-Provence, France
April 11-13, 2002

Tonal Effects on Rhythm in West Middle German

Jörg Peters

Institut für Germanistik
Universität Potsdam, Germany
jpeters@rz.uni-potsdam.de

Abstract

Chanted speech differs from ordinary speech in that its intonational patterns are manifested not only with a characteristic pitch contour, but also with certain rhythmic properties. The chanted call of English and German is a well-known example. It contains a high pitch and a mid pitch, each aligned with a stressed syllable. The pitch of the chanted call is level throughout the syllables that follow the stressed syllables bearing high and mid pitch. When the stressed syllables are closely spaced, they are lengthened. The lengthening provides additional space for sustaining pitch, but it also changes the rhythmic properties of the utterance. The present paper shows that similar effects of tonal structure on rhythm may be found in ordinary speech. We report evidence from three urban vernaculars of West Middle German and propose a phonological analysis that accounts for the similarity of the regional contours to the chanted call.

10.21437/SpeechProsody.2002-127

1. Introduction

In intonational languages like English or German, the temporal alignment of tonal elements is commonly regarded as depending on the rhythmic organization of the text. Autosegmental phonology accounts for this relationship in taking pitch accents as having a single designated tone (the starred tone) that associates with a stressed syllable [1, 2]. Many languages, however, contain a peripheral set of contours that do not simply adapt to the rhythmic structure of the text. These contours share the common feature of being chanted. Take, for example, the chanted call of English and German. The chanted call contains a high pitch and a mid pitch, each aligned with a stressed syllable [3, 4]. The pitch of the chanted call is terraced, i.e. it is level throughout the syllables that follow the stressed syllables bearing high and mid pitch. When the stressed syllables are closely spaced, they tend to be lengthened thus allowing for sustained pitch even if only a small number of syllables is available. As a result, the rhythmic patterning of the utterance changes. According to [5], the chanted call comes with a rhythmic structure that *overrides* the inherent rhythmic structure of the text.

The present paper shows that tonal effects on rhythm may be found in ordinary speech as well. Evidence was found in three urban vernaculars of German spoken in Cologne, Duisburg, and Mannheim. Cologne belongs to the Middle Franconian area. Duisburg is placed about 60 kilometers north of Cologne in the Lower Franconian area. Mannheim is placed about 200 kilometers South of Cologne in the Rhine Franconian area. All three vernaculars pertain to the West Middle German language area. Speakers of these vernaculars use intonation contours that show tonal effects on rhythm similar to those found in the chanted call. This paper examines a preliminary data set and proposes to analyze the contours partly along the lines of the analysis of the chanted call by [5].

2. Speech material

The present study is based on speech data that were recorded from face-to-face conversations among local people, most of them aged between 50 and 75 years. The Cologne data were taken from three episodes of the documentary television series *Die Fußbroichs*. This series, being broadcasted by the *Westdeutscher Rundfunk*, portrays the everyday life of a working class family living in Cologne. The Duisburg data were taken from recordings of eight conversations in the homes of former miners in Duisburg-Homberg and Duisburg-Neumühl (see [6] for details). The Mannheim data were taken from recordings of 11 conversations. Two recordings were carried out by the *Institut für Deutsche Sprache* (Mannheim) and nine by the research project *Dialektintonation* (Potsdam and Freiburg). About 700 utterances were selected for acoustical analysis from 28 hours of recorded speech. F_0 traces were computed with the help of the analysis program PRAAT (© 1992-2001 by Paul Boersma and David Weenink).

3. Cologne Data

In Cologne German, many questions bear a pitch contour that consists of a mid pitch on the nuclear syllable, a local peak after the nuclear syllable, and a final fall at the end of the intonational phrase (IP). The postnuclear peak aligns with the head of the last foot in the IP. If this position is occupied by the nuclear syllable, the peak aligns with the last syllable of the IP. As feet are left-headed in Cologne German, this syllable is a weak syllable unless the last foot is monosyllabic. The final fall is only fully realized if sufficient sonorant material is available.

In some cases, the nuclear syllable is lengthened and the pitch on this syllable is level. Lengthening and level pitch are most likely when the utterance meets two conditions: (i) the nuclear stress is located on the head of the last foot, (ii) the nuclear syllable contains a long vowel, a diphthong, or a short vowel and a sonorant that is not ambisyllabic. If (i) holds and the last foot is bisyllabic, the nuclear syllable goes on the penultimate syllable and the peak aligns with the ultimate. In this case, the last syllable tends to resist reduction processes that usually take place in other tonal contexts. This is particularly true of syllables containing a schwa and a sonorant like [əl], [əm], and [ən] which normally tend to be reduced to [l̩], [m̩], and [n̩] or to be attached to the preceding syllable after deletion of schwa.

Figure 1 illustrates level pitch on a nuclear syllable with a long vowel. As can be seen from the F_0 trace, the pitch plateau spans the nucleus (*y*) rather than the whole syllable.

This can also be observed in syllables containing a diphthong. In syllables with a short (lax) vowel and a sonorant coda, lengthening applies to the coda without neutralizing the distinction between tense and lax vowels. If the nuclear syllable contains a short (lax) vowel and a non-sonorant coda, lengthening of the non-sonorant coda may occur but the pitch on the nuclear syllable is not level as the lengthening does not apply to a sonorant part of the syllable.

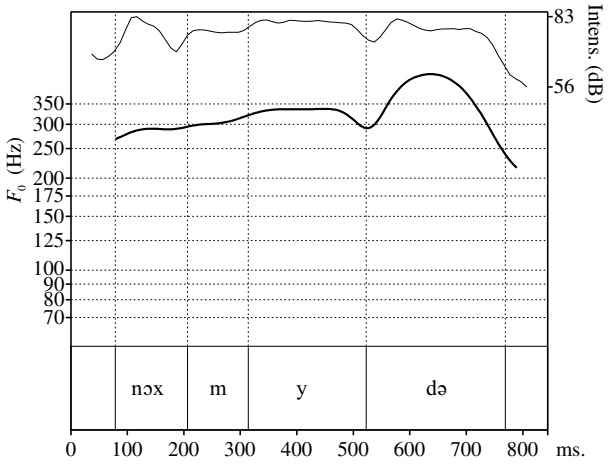


Figure 1: F_0 trace of the utterance *Noch müde? (Still tired?)*

If the nuclear stress is located on the last syllable of the IP, lengthening and level pitch may likewise occur. However, the plateau formation is restricted to the first part of the lengthened syllable and followed by a local peak on the second part. As a result, the rise to the final peak is perceived as stepping rather than gliding.

Finally, lengthening and level pitch may also be observed in nuclear syllables that are located on the penultimate foot, provided this foot is monosyllabic. In this case, the final peak aligns with the next syllable, as this is the last stressed syllable of the IP.

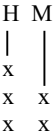
In conclusion, lengthening of the nuclear syllable and level pitch is most likely if either the nuclear syllable is followed by a syllable that attracts the final peak or the nuclear syllable is in IP-final position.

4. Phonological Analysis

Our analysis rests upon the assumption that both the sustained pitch and the durational features of the Cologne contour are structurally related to corresponding features of the chanted call. We assume that the peculiarities of both contours can be explained by using partly the same principles, i.e. *Beat Splitting* and *Tone Sharing* as introduced by Hayes and Lahiri [5].

The English calling contour consists of a sequence of three pitches, low, high, and mid. The low pitch is optional and the mid pitch is said to be one minor third lower than the high pitch [3, 7]. The English calling contour comes in two versions, one chanted and one not. According to [5], the chanted version differs from the non-chanted version both by terracing of pitch and by syllable lengthening, depending on the number of syllables available. The lengthening may even override distinctions of segmental length. Syllable lengthening affects the initial syllable of a tone domain (defined by [5] as the sequence of syllables on which a given tone is realized). If the tone domain consists of a single syllable, lengthening is obligatory. If the tone domain consists of two syllables, lengthening is optional. If it consists of three syllables, lengthening is dispreferred. If it consists of four or more syllables, it is impossible.

To account for the tonal facts of the chanted call, Hayes and Lahiri assume that the chanted call comes with a metrical grid. The underlying representation of the chanted call is (1):

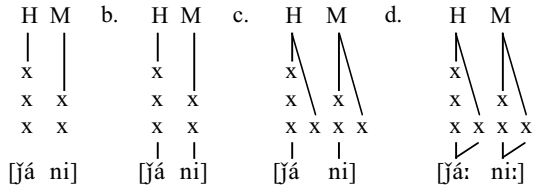
(1)


The strong beat bearing the H tone links to the syllable that bears the main stress (the nuclear syllable). The strong beat bearing the M tone links to the strongest stress following the H tone. If no stressed syllable comes after the main stress, the strong beat bearing M associates with the final syllable.

According to [5], the rhythmic structure of the chanted call has to meet a constraint called *Obligatory Offbeat Condition* (OOC) (see also [8]):

(2) Any strong beat must be directly followed by a weak beat.

The grid of the chanted call in (1) *violates* the OOC. In order to be instantiated in actual speech, additional beats must be inserted after the strong beats. This process is called *Beat Splitting*. An additional principle, called *Tone Sharing*, links the tones to the additional beats. (3) illustrates the application of Beat Splitting and Tone Sharing by the example *Johnny* taken from [5]:

(3)


(3a) gives the underlying representation of the chanted call. In (3b), the strong beat bearing the H tone links to the main-stressed syllable of the text. The strong beat bearing the M tone links by default to the final syllable. (3b) violates the OOC. Hence, Beat Splitting adds a weak beat after each strong beat in (3c). In addition, both the H tone and the M tone are linked to the inserted beats by Tone Sharing. In (3d), the inserted beats associate with the next syllable to the left. As a result, these syllables are lengthened.

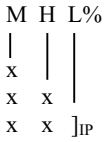
To account for the tonal and durational facts of the Cologne contour, we likewise assume a tonal sequence that is linked to a metrical grid. However, we must bear in mind that the Cologne contour differs from the English chanted call in at least three respects:

(i) *Tonal structure.* The Cologne contour consists of a mid pitch followed by a high pitch whereas the English chanted call consists of a high pitch followed by a mid pitch. In addition, the Cologne contour ends with a final fall whereas the final boundary of the chanted call is tonally unspecified. We characterize the Cologne contour by M* H-L%. M* accounts for the fact that the nuclear syllable is realized at mid-level rather than at a high or low level. Furthermore, the nuclear syllable does not seem to be lowered or raised due to additional emphasis, i.e. M does not appear to be "stress-sensi-

tive". H- is a phrasal tone that we take to be secondarily associated with the head of the last foot, in the sense of [9, 10]. If the nuclear stress is located on the stressed syllable of the last foot, the phrasal tone is secondarily associated with the last syllable of the IP. Finally, L% accounts for the fact that the contour ends with a final drop to the baseline if enough sonorant material is available after the last peak.

For present purposes, we assume that (4) is the underlying form of the Cologne contour:

(4)



Analogously to the analysis of the chanted call by [5], we hypothesize that the Cologne contour is underlyingly linked to a schematic grid. However, the order of the H tone and the M tone is reversed. Furthermore, we stipulate an additional tone that is linked to the final boundary of the IP.

(ii) *Limited occurrence of level pitch.* In the English chanted call, pitch is level throughout the syllables that follow the stressed syllables bearing high and mid pitch. In contrast, the level pitch of the Cologne contour is limited to the nuclear syllable. The rhythmic structure of the Cologne seems to meet a weaker constraint than the OCP. We assume that it complies with a *Clash-Avoidance Provision* (CAP) as stated in [5]:

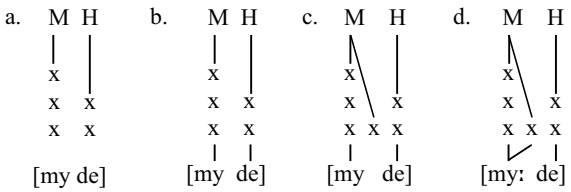
(5) Strong beats should not be adjacent.

(4) violates the CAP. We assume that this conflict is solved by Beat Splitting and Tone Sharing as in the analysis of the chanted call. There is no need, however, that Beat Splitting applies to both strong beats that bear a tone. Inserting an additional beat after the beat bearing the M tone will suffice to make (4) consistent with the CAP. The exclusive application of Beat Splitting to the beat that bears the M tone accounts for the fact that in the Cologne contour lengthening and plateau formation are limited to the nuclear syllable.

(iii) *Relation between plateau formation and lengthening.* In the English chanted call, pitch is level even in the absence of syllable lengthening. In the Cologne contour, it is level only if the nuclear syllable is lengthened. Thus, the rhythmic structure of the Cologne contour does not simply override the inherent rhythmic structure of the text. It seems more likely that there is a matching between the inherent rhythmic structures of tune and text. The first overrides the second only if both are not compatible. Beat Splitting and Tone Sharing do only apply if otherwise the strong beat bearing the M tone would be followed by the strong beat bearing the H tone. In addition, if only an unstressed syllable is available as a docking point for the strong beat that bears the H tone, the association of tune and text makes it strong.

We illustrate this line of analysis by the utterance *Noch müde?* (cf. Figure 1; the boundary tone L% is ignored):

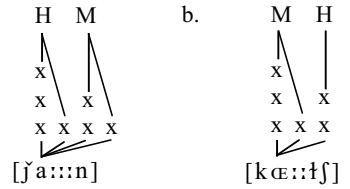
(6)



(6a) gives the underlying form of the Cologne contour. In (6b), the strong beat bearing the M tone is linked to the main stress. The strong beat bearing the H tone is linked to the final syllable. As (6b) violates the CAP, Beat Splitting and Tone Sharing apply in (6c). However, only the beat that is linked to the M tone is split. There is no need to split the beat linked to the H tone. Similarly, Tone Sharing applies to the M tone but not to the H tone. This accounts for the fact that, in Figure 1, we found a plateau on the nuclear syllable but not on the last syllable. Finally, the newly inserted weak beat is linked to the nuclear syllable (6d). As a result, the nuclear syllable is lengthened. In addition, the linking of the strong beat bearing the H tone with the (weak) syllable *de* accounts for the fact that this syllable resists reduction, which may be expected in other tonal contexts.

The previous analysis of the chanted call was restricted to utterances that contain at least one additional syllable after the main stress. If the main stress is located on the final syllable of the IP, both the strong beat bearing the H tone and the strong beat bearing the M tone associate with the same syllable. In this case, the single syllable is "broken" into a number of metrical positions (cf. [3]) and undergoes additional lengthening. As a result, the last syllable is realized both with a high and with a mid level pitch like in the example of *John* (7a).

(7)



A similar analysis applies to the Cologne contour. If the main stress is located on the final syllable, Beat Splitting and Tone Sharing likewise apply. However, Beat Splitting applies only to the strong beat bearing the M tone. Accordingly, the stressed syllable is realized with two pitches, but only the mid pitch is level. This is illustrated by the example *Noch immer Kölsch?*, '*still Kölsch* (brand of beer)*?*', in (7b), with the main stress on *Kölsch*. According to (7b), *Kölsch* is associated with three beats, two of which are associated with the M tone. This, together with the fact that both tones are linked to the text over strong beats, may account for the observation reported in § 3 that the Cologne contour is perceived as stepping rather than gliding when the nuclear syllable occurs in IP-final position.

5. More data

In Cologne German, plateau formation is not restricted to question intonation. A rather similar contour was found in utterances by which speakers signal turn continuation. It resembles the question contour except that the final pitch does not drop to the baseline. In most cases, there is only a small drop, which may be due to declination rather than to the presence of a final boundary tone. Accordingly, we characterize this contour by M* H-%, with "%" for a tonally unspecified IP boundary (see [11] for a similar notation).

Furthermore, Cologne German does not seem to be unique in having contours with level pitch in ordinary speech. Evidence comes from Duisburg and Mannheim. We found speakers of both cities using similar contours in similar situations, i.e. in asking a question and in signaling turn continuation. All contours consist of a nuclear M (or L) tone followed

by a high phrasal tone that is linked to the last stress of the IP. When the nuclear syllable is located in the last foot, the phrase tone uniformly goes to the last syllable of the IP. Tables 1 and 2 summarize the main facts (see [12] for a more detailed analysis).

Table 1. Question intonation.

		Nuclear syllable	
	Tune	in non-final foot	in final foot
Köln	M* H-L%	Last stress	Final syllable
Duisburg	M* H-%	(No data)	Final syllable
Mannheim	L* H-L%	Last stress	Final syllable

Table 2. Contours signaling turn continuation.

		Nuclear syllable	
	Tune	in non-final foot	in final foot
Köln	M* H-%	Last stress	Final syllable
Duisburg	M* H-%	Last stress/final	Final syllable
Mannheim	M* H-L%	Final syllable	Final syllable

All contours listed in Tables 1 and 2 tend to come with syllable lengthening and level pitch under the conditions given in § 3 for the Cologne question contour. The nuclear tone (M or L) and the H phrase tone must be linked to adjacent syllables and the nuclear syllable must be "stretch-able" (it must contain a long vowel, a diphthong, or both a short vowel and a sonorant that is not ambisyllabic). These facts suggest that all six contours can be accounted for by using the principles proposed in the analysis of § 4.

6. Discussion

In ordinary speech data from Cologne, Duisburg, and Mannheim, we found intonational patterns that are manifested not only with pitch, but also with certain rhythmic properties. The pitch contours consist of a nuclear M (or L) tone, a high phrasal (H-) tone and (optionally) a final low boundary tone (L%). The nuclear syllable is lengthened and bears level pitch when it occurs on the penultimate syllable. When it occurs on the final syllable, it undergoes additionally lengthening and attracts both the nuclear M (or L) tone and the phrasal H tone. In this case, the nuclear syllable bears both a short level pitch and a final peak.

In the present paper, we argued that the German contours could be explained by using partly the same principles as introduced by [5] in the analysis of the English chanted call. In particular, we attributed the lengthening of the nuclear syllable and the accompanying level pitch to the application of Beat Splitting and Tone Sharing. Thus, even if the contours discussed in this paper are clearly not perceived as being chanted, our analysis suggests that, from a phonological point of view, the difference between chanted and non-chanted speech is possibly not as clear-cut as it may appear at first sight.

Finally, all German contours that attest the operation of Beat Splitting and Tone Sharing in ordinary speech belong to varieties of West Middle German. No evidence was found in other language areas including Northwest Germany (Hamburg), Southwest Germany (Freiburg), and East Germany (Berlin, Dresden). A further investigation of the structure and geographical distribution of the contours presented here may therefore be of particular importance for establishing a typology of German intonation.

7. References

[1] Pierrehumbert, J. B. 1980. *The phonology and phonetics of English intonation*. PhD thesis, Massachusetts Institute of Technology. Indiana University Linguistics Club.

[2] Ladd, D. R. 1996. *Intonational Phonology.* Cambridge: CUP.

[3] Liberman, M. 1975. *The intonational system of English*. Diss. 1975. New York: Garland, 1979.

[4] Gibbon, D. 1976. *Perspectives of intonation analysis*. Frankfurt a. M.: Lang.

[5] Hayes, B.; Lahiri, A., 1992. Durationally specified intonation in English and Bengali. In *Conference on Music, Language, Speech, and Brain. Proceedings of the 1990 Wenner-Gren Center,* R. Carlson; L. Nord; J. Sundberg (eds.), 78-91.

[6] Salewski, K., 1998. *Zur Homogenität des Substandards älterer Bergleute im Ruhrgebiet.* Stuttgart: Steiner.

[7] Leben, W. R. 1976. The tones in English intonation. *Linguistic Inquiry* 21, 69-107.

[8] Selkirk, E. 1984. *Phonology and syntax: The relation between sound and structure.* Cambridge, Mass.: MIT Press.

[9] Pierrehumbert, J. B. & Beckman, M. E. 1988. *Japanese tone structure*. Cambridge, Mass.: MIT Press.

[10] Grice, M.; Ladd, D. R.; Arvaniti, A., 2000. On the place of phrase accents in intonational phonology. *Phonology* 17, 143-185.

[11] Grabe, E. 1998. *Comparative intonational phonology: English and German*. Wageningen. Diss. University of Nijmegen.

[12] Peters, J. (in prep.). *Regionale Variation der Intonation im Deutschen*. Habilitationsschrift, University of Potsdam.

Acknowledgements

The research reported here is part of the project *Dialektintonation* supported by the *German Research Foundation* (DFG) under project No. SE 699/3-3. I would like to thank Margret Selting (University of Potsdam), Peter Auer, Peter Gilles (both University of Freiburg), and Margit Aufterbeck (University of Cambridge) for their comments. I am also grateful to Arend Mihm (University of Duisburg) and Werner Kallmeyer (*Institut für deutsche Sprache*) for making available audio recordings from their own research projects.