Acoustic Patterns and Communicative Functions of Phrase-Final F0 Rises in German: Activating and Restricting Contours

Ernst Dombrowski
Institute of Psychology, Christian-Albrecht-University, Kiel, Germany

Oliver Niebuhr
Institute of Phonetics and Digital Speech Processing (IPDS), Christian-Albrecht-University, Kiel, Germany

Short title:
Activating and Restricting Contours

Postal address:
Ernst Dombrowski
Lindenstrasse 13
24118 Kiel
Germany

Phone: +49 431 568328
e-Mail: ed@ipds.uni-kiel.de

Abstract
Acoustic features and communicative functions of phrase-final F0 rises starting before an accented-vowel onset are analysed in a corpus of German unscripted speech. Two conversational conditions are examined: turn-yielding and turn-holding. The most important feature distinguishing rises in these two conditions is the *range proportion*, which differentiates between two patterns as follows: (1) raised pitch on the accented syllable and restrained pitch movement in the tail of the contour, (2) lowered pitch on the accented syllable and extended pitch movement in the tail. The first pattern is seen as a *restrictive* gesture, e.g. preventing the dialogue partner from turn taking. The second one is viewed as an *activating* gesture, inviting the co-participant to contribute.

1 Introduction

1.1 Preliminary remark

This contribution reports a function-based acoustic analysis of phrase-final rising intonation in German, focusing on a particular contour type. Accent-related pitch patterns are dealt with that already rise into the accented vowel and continue rising up to the end of the prosodic phrase. The phonetic variation of these patterns is examined in two conversational settings, turn-holding and turn-yielding, with regard to aspects of form and function in rising intonation which have been missed in previous research -- or which have been treated less systematically. On the one hand, a novel perspective for the description of phrase-final rising is suggested; on the other hand, the communicative features of rising contours are reconsidered.

1.2 Phonological basis

As a phonological basis underlying the analysis, an approach to intonation is used which conceptualizes accent-related patterns as integral units, describing them in terms of contours instead of tonal sequences. Consequently phrase-final rises will be characterized by a set of features that capture their pitch course as a whole. For German, such an approach is represented in the *Kiel Intonation Model, KIM* [Kohler, 1991a, b, 1997].

In KIM, the elements of intonation are *accent-related peak and valley contours*. They derive from two basic forms which are not decomposed: rising-falling and falling-rising. At least three peak categories and two valley categories are distinguished for German according to the position of their F0 maximum or minimum, respectively -- relative to the onset of the accented vowel. There are early, medial, and late peaks; and there are early and late (i.e., non-early) valleys [Kohler, 1991a]. Both basic contour shapes of KIM, peaks and valleys, have been subjected to systematic investigation concerning their perceptual relevance, phonological form, and communicative function [Kohler, 1987, 1991c; Gartenberg and Panzlaff-Reuter, 1991; Landgraf, 2003; Niebuhr, 2003; Niebuhr and Kohler, 2004]. Based on KIM and the labelling tool PROLAB [Kohler, 1995; Kohler, Pätzold, and Simpson, 1995; Peters and Kohler 2004], large corpora of German speech have been prosodically annotated (Kiel Corpus of Read and Spontaneous Speech [IPDS, 1994, 1995, 1996, 1997]).

Phrase-final rises, which will be examined here, are described in terms of *valley contours*. For the distinction of the different types of valleys in KIM, the position of the F0 minimum and, in particular, the onset of the rising movement are important. Early valleys have their F0 minimum before the accented-vowel onset and rise already into the accented vowel. Late valleys have the F0 minimum and the rise after the accented-vowel onset [Kohler, 1991a;

Peters and Kohler, 2004]. Another feature of (phrase-final) valley contours in KIM is their height [Kohler, 1991a; Peters and Kohler, 2004].

1.3 Domain of investigation

The present investigation is in line with previous perceptual and functional studies carried out in the framework of the Kiel Intonation Model. It has been prompted by informal observation comparing turn-internal and turn-final rising intonation in dialogue data from the Kiel Corpus [IPDS, 1995, 1996, 1997]. Phonologically, this study is limited to just one of the two valley categories of KIM, viz. the *early valley*, which is examined in phrase-final position. The contours allocated to this category show considerable phonetic variation, which points to a further formal and functional distinction that has not been investigated up to now. This variation is systematically analysed in a modelling of valley contours and a tentative communicative explanation.

1.4 Examples

The phonetic and functional differences observed in phrase-final rises of the early valley type are illustrated with two examples (Fig. 1a, b). The examples have been taken from dialogue data elicited in an appointment-scheduling scenario [Kiel Corpus of Spontaneous Speech, Vol. II, IPDS, 1996: g315a009 and g311a005]. The participants of the dialogues had to arrange business meetings, working sessions etc. They sat in separate recording rooms and had a telephone-style conversation via headset. The recording procedure required the speakers to press a button as long as they were speaking. They released the button when they had finished their turn. While a speaker was pressing the button the channel was not free for the other dialogue partner [for further details, see Kohler et al., 1995]. Thus, two conversational conditions can be unambiguously identified within the recordings: *turn-yielding*, which is always *non-competitive* in this frame, and *turn-holding* [cf. Selting, 1995]. For both examples a description of the contour is given and a functional paraphrase is added. Sound files and a more detailed transcription of the texts are available in a separate html document [Dombrowski and Niebuhr, 2005].

/////////INSERT FIGURE 1A AND B ABOUT HERE/////////

Example (a) (... *Osters'onntag und Osterm'ontag*; „... Easter Sunday and Easter Monday"; with the accented vowel marked by an apostrophe). In example (a), the phrase-final rise on *Osterm'ontag* is preceded by a slightly falling course across a stretch of several syllables which leads to an F0 minimum on the pre-accentual syllable *ter*. F0 remains low in the onset of the accented syllable *m'on*. The rise begins with a small but clear step towards the accented vowel. Extended further rising follows, accelerating towards the end of the prosodic phrase. The interval between the F0 minimum and F0 at the end of the accented syllable is ca. 2 semitones in this example. The further rising covers ca. 8 semitones.

In this example the speaker suggests the date for an appointment within a series of appointments. He signals to the other participant of the dialogue that he may agree or make another suggestion (thus encouraging him to make his own choice). Then he passes the turn. The phrase presented here depends grammatically on a preceding unit with declarative syntax.

Example (b) (... *Anfang Dez'ember*; „... at the beginning of December"). The second example is by the same speaker. At the end of the contour stretch before the rise, pitch is lowered (as in example (a)). From there, the accented vowel is reached in a medium upward step. The nucleus and the rhyme of the accented syllable *z'em* show clear further rising. Then the upward movement is restrained. There is a small step towards the vowel of the unaccented

syllable, and on this vowel, there is only slight rising. In example (b) the interval towards the accented vowel is ca. 5.5 semitones; the further rise covers ca. 2 semitones.
Again the speaker suggests the date for an appointment. But here he fixes the date more precisely in the following phrases. Thus, the non-speaking participant in the dialogue is given a signal to continue listening and not to speak until he has received the message completely (the other participant is restricted in his conversational alternatives). The phrase also depends grammatically on a preceding one with declarative syntax. Further examples with tentative functional interpretations are given in Dombrowski and Niebuhr [2005].

Starting from the above characterization, the examples (a) and (b) can be compared with respect to both their *phonetic substance* and their *functional background*.

1.5 Phonetic substance
Three basic dimensions may be included in the phonetic description of phrase-final rises and may thus be checked for the comparison of contours. The three dimensions apply to both the early and late valleys according to KIM. They comprise (1) the *location* of the initial F0 minimum before the rise and, in particular, of the rising movement itself, (2) the *range* of the rise and (3) its *shape*, i.e., its curvature. The dimensions can be characterized by their polar expressions -- as shown in Figure 2. They will be applied to the two examples.

/////////INSERT FIGURE 2 ABOUT HERE/////////

The location of the rise: The position of the rising movement is the same in the examples (Figures 1(a) and (b)). The contours have their F0 minimum before the accented-vowel onset and rise into the vowel (thus, they are *early valleys*). As, in the present approach, there is no serial decomposition of accent-related contour shapes implementing individual tones as basic units, the comparison of the patterns in the two examples can do without further points of reference in the temporal domain at this stage.

The range of the rise: Example (a) has a higher rise than example (b). Consequently, the contours differ in their total range. This difference can be interpreted in terms of an attribute of the entire pitch course. But it may also point to a particular role of the end of the contour.

The shape of the rise: The most salient feature distinguishing the contours in the examples is their curvature. Both patterns rise continuously starting with an F0 minimum before the accented-vowel onset. But the step into the rising pattern is small in the one case and pronounced in the other, compared with the total range of the contour; in turn, the further rising is extended or reduced. In consequence, the rises may be described on a convex-concave scale with one being curved inwards and the other outwards, which is a feature of the contour as a whole.

Summing up, the above instances of phrase-final early valleys are different in two out of three dimensions within the descriptive template: They differ in their range and their contour shape. Both range and shape can be conceived of as continuous variables. The range, or height of the end of the contour, can be raised or lowered. The shape may be varied by moving a significant point in the middle of the contour up-- or downwards. This pivot would be a perceptually relevant event; in early valleys this can be represented as the perceived step from the pre-accentual pitch minimum into the accented stretch of the contour.

1.6 Functional background

The contours in examples (a) and (b) seem to be related to a communicative distinction that is relevant to particular conversational situations. The organization of the dialogues by an external criterion (pressing a button vs. releasing a button) points to a contrast between *turn-holding* and voluntary *turn-yielding*. The (tentative) functional interpretations of the dialogue turns add further aspects. As an underlying feature, the contours may be used by speakers to control the behaviour of the other participants in a conversation indicating that they do or do not have alternatives. This aspect is emphasized in the interpretations of further examples in Dombrowski and Niebuhr [2005]. In a statistical sense, the pattern in example (a) may often be associated with questions. Questions, in turn, often occur together with turn-yielding. However, there is no deterministic relation between sentence types, speech acts, and the two patterns of examples (a) and (b). E.g., both utterances in the examples depend on preceding declarative syntax. The non-deterministic relation between syntax and contour type is again backed up by materials in Dombrowski and Niebuhr [2005]. The further analysis of form and function in rising contours requires a clear definition of the communicative conditions which are assumed to trigger the use of different patterns. The more clear-cut these functional classes are, the more accurate can be the description of the related acoustic patterns.

1.7 Questions and hypotheses

The comparison of the above examples has demonstrated that the early valley category of KIM can show systematic phonetic variation. The variation is significant, because it may be associated with a distinction in communicative function; this relationship between substance and function is also supported by Dombrowski and Niebuhr [2005]. The question then arises as to whether this evidence from individual observations is representative of larger sets of conversational data. To answer this question a comprehensive acoustic analysis is carried out based on the entire dialogue materials that the above examples have been taken from [Kiel Corpus of Spontaneous Speech; IPDS, 1995, 1996, 1997]. The objective of the analysis is to measure early valley contours in two well-defined conversational conditions forming different classes -- in accordance with the above case observations. The results may contribute to understanding the role of rising patterns in interpersonal interaction and may at the same time provide evidence concerning their formal description.

Defining the functional classes for the analysis benefits from the recording procedure used in the appointment-scheduling dialogues, which provides clear external criteria of (non-competitive) turn-yielding vs. turn-holding. Thus, an important conversational feature presumably contributing to phonetic variation within phrase-final early valleys is controlled by design.

The contours allocated to the two conversational conditions are analysed in a multivariate perspective starting with the three-dimensional template of position, range and shape of the rise. Measures of shape and measures of range are included. In this, variables can be accounted for which refer to accent-related patterns as a whole or to sub-sections of these patterns. Duration features will be used additionally to get a full representation of the course of the contours under investigation.

In a first step the differences between the two functional classes are tested for the individual acoustic variables. Next, the contribution of the variables to the distinction of patterns will be checked. Finally, the prediction of the functional classes by the included variables is calculated. Two clearly distinct sub-patterns are expected within the early valley category that particularly differ in acoustic parameters characterizing the overall shape of the contour.

2 Method

2.1 Contour model and setting of measurements

A contour model is created from the combination of the three phonetic dimensions location, range, and shape, proposed as a template to describe phrase final rises (cf. Fig. 2). The range dimension concerns the frequency interval between the onset and the end of the rise. A change in the shape of a continuous rise means a variation along a continuum from concave to convex. This can be modelled conjointly by implementing pivots in the rising movement dividing it into sections. Each of the sections can vary in the time and frequency intervals, in this way contributing to the phonetic dimensions shape and range. Finally, the location dimension determines the position of the onset of the rising movement.

This general phonetic framework can be adapted to different categories of phrase-final rises, e.g. by changing the number of pivots, their temporal positions and the starting-point of the rising movement. In the present study, it is adapted to the early valley category. In this, the location dimension is restricted to rises starting from a minimum before the accented-vowel onset. Furthermore, the number of pivots is set to one. From an acoustic point of view, a single pivot is regarded sufficient to approximate the variation in shape of the F0 contours observed in the two conversational settings turn-holding and turn-yielding (cf. the examples in Fig. 1a and b). This acoustic conceptualization is combined with the authors' perceptual impression that early valleys have a single centre of gravity within the rising movement. The height of the perceived centre within the pitch range of the contour follows the acoustic variation from concave to convex. An examination of examples shows for polysyllabic accent groups that the height of the perceived centre can best be approximated by the highest F0 value within the accented syllable. Since the analysis deals with continuous rises, this value is usually located at the end of the voiced part of the accented syllable. For monosyllabic accent groups it is much harder to determine an acoustic equivalent for the perceived centre of the rise. These accent groups are therefore disregarded in the analysis, which indicates that the current model for early valleys is only provisional. But also for the polysyllabic accent groups there is no experimental evidence that the height of the pivot and its representation by the highest F0 value of the accented syllable is independent of variable durations of the rising movements, e.g. due to different speaking rates or a different number of (unaccented) syllables between the accented syllable and the end of the prosodic phrase. Finally, regarding these assumptions, it must be pointed out that the pivot is seen as a phonetic tool to make changes in the shape of the rising movement measureable and must not be confused with a phonological target anchored in the segmental string.

/////////INSERT FIGURE 3 ABOUT HERE//////////

The model of the early valley is illustrated in Figure 3. The pivot divides the rising movement into two sections. The frequency interval of the first section, *range 1*, covers the distance from the prevocalic minimum of the F0 rise to the highest F0 value within the accented syllable. *Range 2* represents the remaining frequency interval from the highest F0 value within the accented syllable to the end of the voicing before the prosodic phrase boundary. The corresponding time intervals, *duration 1* and *duration 2*, extend from the minimum of the F0 rise to the end of the accented syllable and from this point to the end of voicing in the phrase, respectively. Hence, unvoiced events between the end of voicing and the phrase boundary are disregarded in the time and frequency measurements.

In order to characterize the rising contour as a whole, we combined the frequency values as well as the duration values in two different ways, obtaining absolute and relative measurements. By summing *range 1* and *range 2*, the *total range* was determined. In the same way, the *total duration* was created by summing *duration 1* and *duration 2*. With respect to the relative measurements, the *range proportion* was obtained by dividing *range 1* by the *total range* (*range 1/[range 1 + range 2]*). This proportional value indicates the position of the pivot of the contour in relation to the overall range. It can take values from 0 to 1. Analogously, the *duration proportion* was calculated by dividing *duration 1* by the *total duration* (*duration 1/[duration 1 + duration 2]*). The absolute and relative types of measurements are mathematically independent of each other.

Altogether, eight acoustic features were taken into account: the original four frequency and duration distances (*range 1, range 2, duration 1, duration 2*) and four variables formed by combinations of these distances, namely *total range, range proportion, total duration* and *duration proportion*. In these variables pitch and time are treated independently, i.e. the eight features do not combine time and F0 information. Another way to capture the quality of a rising movement would be to calculate the *slopes* of the contour between the three points defined and include them in an index of the contour shape. However, in the calculation of contour slopes the duration of the syllables and segments must be accounted for. The variable syllabic and segmental complexity of the contours would increase the error variances; and thus, a possible distinction between the patterns found in the different functional environments would become less clear. The effect of the syllabic and segmental structure on the slope of contours can be seen from Grabe's [1998] finding that German (and English) show compression in the phrase-final rises. As a conclusion for the present investigation, a description of rising contours should be preferred that is uninfluenced by the number of syllables or segments. The relative values of the variable *range proportion* meet this requirement. As a control, a comparable index of the shape was calculated based on slopes, the *slope proportion* (*[range 1/duration 1]/[total range/total duration]*).

Finally, every analysis involving F0 measurements has to address the problem of microprosodic perturbations of the F0 course (i.e. intrinsic F0) due to aerodynamic and muscular interactions with consonants and vowels, as they are for example described by Gartenberg and Panzlaff-Reuter [1991] and Lehiste [1970]. Since with regard to recent research in this field, neither the context-dependent amount of intrinsic F0 nor its perceptual relevance seems to be fully understood and thus fully predictable [cf. Fowler and Brown, 1997; Niebuhr, 2004] it was considered impossible to exclude this factor from the analyis. Hence, it was not accounted for in the measurements. It was assumed, however, that pooling the data for every F0 point (e.g. the F0 minimum) amalgamates measurements in various segmental environments, thus creating an intrinsic noise, which does not influence the data in a systematic way and therefore can be neglected. Nevertheless, these influences should be kept in mind.

2.2 Formation of the samples

A symbolic search in the Kiel Corpus of Spontaneous Speech [IPDS, 1995, 1996, 1997], using the KIM-based PROLAB labels [cf. Kohler, 1995; Kohler et al., 1995; Peters and Kohler, 2004], resulted in a database of about 500 early valleys. These were further filtered and sorted to arrive at two samples of contours, assigned to the turn-holding and turn-yielding conditions.

First, early valleys had to be excluded from the analysis that did not fit into the contour model described in 2.1. This concerned all valleys produced in monosyllables. Further, since the onset of the rise of the early valley had to be clearly identified, all valleys were omitted that followed other peak or valley contours without a noticeable indentation in the F0 course. Moreover, phrase-final early valleys preceded by other valley contours often start already very high within the modal frequency range of the speaker. The small frequency range remaining for the phrase-final early valleys suggests that they have to be treated differently from early valleys following peak contours. So, we disregarded early valleys preceded by other early or late valleys in the analysis, regardless of whether such a sequence was integrated into a single rising movement or concatenated by indentation. Finally, for each of the remaining cases the prosodic labelling was perceptually checked and mistakes were eliminated. At the end of this filtering cascade, 177 relevant early valleys remained. Their time and frequency measurements were taken according to the model presented in 2.1. The measurements were performed using the speech analysis software xassp (see www.ipds.uni-kiel.de/forschung/xassp.en.html) developed at the IPDS. Obvious mistakes of the F0 analysis like octave errors were corrected manually.

After this, the 177 early valleys were assigned to either the turn-yielding or the turn-holding condition. This could be done reliably with reference to the special feature of the recording procedure used in the corpus to control the interaction of the dialogue partners. Using the speech-external actions of button pressing/releasing as the decisive criterion (cf. 1.4), the assignment was independent of syntactic, semantic or prosodic features or interpretations. Consequently, all early valleys immediately preceding the action of releasing the button constituted the turn-yielding sample.

Compared with this, phrase-final early valleys occurring while the speaker kept the button pressed, thus giving an explicit speech-external signal that s/he wanted to continue, were candidates for the turn-holding sample. However, it had to be considered that among these candidates there were either phrases showing an interrogative syntax or depending on phrases with interrogative syntax. In communication, the characteristic use of interrogative syntax is to signal to the dialogue partner that s/he is supposed to contribute to the conversation. Thus, interrogative syntax has a certain implication of turn-yielding, and the functional status of such phrases in turn-internal position is therefore not clear with regard to turn-holding and turn-yielding. So, it is problematic in this position to include phrase-final early valleys from interrogative contexts in the turn-holding sample together with early valleys from other syntactic constructions. Consequently, within the turn-internal environment, the data set was restricted to early valleys in phrases that had a declarative syntax or depended on such phrases. This also meant leaving out early valleys in turns that only contained phrases with elliptical syntax since these could neither be identified as declarative nor interrogative by means other than intonation.
So finally, 106 early valleys were assigned to turn-yielding or turn-holding. They came from 61 dialogues, involving 8 female and 18 male speakers. The turn-yielding sample consisted of 32 early valleys. Of these, 10 occurred in turn-final ellipses. Wh-Questions as well as yes-no questions contributed 7 valleys each. Another 8 valleys came from clauses with declarative syntax. The turn-holding sample, on the other hand, comprised 74 early valleys, 34 of which were located at the end of declarative main clauses preceding other declarative main clauses. Another 19 cases came from prosodic phrases within declarative main clauses. The remaining 21 valleys were either located at the end of declarative main clauses followed by other incomplete declarative clauses (13) or at the end of the first part of complex sentences (8). Further, both samples were comparable with regard to the number of syllables between the accented one and the end of the phrase. The number varied between 1 and 5, with most of the cases showing either one or two unaccented syllables. In the turn-yielding sample, there were 23 instances (72%) with 1 unaccented syllable, another 6 (19%) with 2 and 1 (3%) with 3 unaccented syllables. The corresponding distribution in the turn-holding sample was 58 (78%), 13 (18%) and 2 (3%).

2.3 Methods of data analysis

The statistical analysis of the corpus measurements was done in three steps. At first mean differences were tested for the two conversational conditions (turn-yielding vs. turn-holding) and for the syntactic patterns occurring within the two conditions. Eight dependent variables (acoustic features) were included employing F tests and single *t* tests for each variable. Second, correlations were calculated between all variables. Thirdly, two discriminant analyses were run, based on different predictor sets [Tatsuoka, 1971]. These analyses yielded (a) a multivariate overall comparison of the two conversational conditions, (b) a test of the contribution of each predictor to the discrimination of the conditions, (c) a test of the predictive power of all variables included. The statistical analysis was done using SPSS [Brosius, 2002].

3 Results

3.1 Tests of mean differences for eight variables in the two conversational conditions
Means, standard deviations, and probabilities for the eight acoustic features examined are summarized in Table 1. The tests of mean differences reveal a clear contrast between the two conversational conditions turn-yielding and turn-holding. The differences concentrate on the variables involving pitch. The strongest effects are found for the *range proportion* whereas no significant differences show up for the *total range*. However, a separate look at *range 1* and *range 2* reveals that the *total range* is composed of two contrasting aspects both contributing to the distinction of the two conversational conditions. Excluding the data from two pairs of male speakers (13 dialogues) the mean differences of the total range increase to 8.6 vs. 10.3 semitones in the turn-holding vs. turn-yielding (cf. 3.4 for a justification of this exclusion). These mean values are in accord with ranges reported for German question and continuation contours, e.g., Adriaens [1990]. A further difference was found for one of the partial durations (*duration 1*) which is shorter in the turn-yielding condition. As a control, a slope-based index of the contour shape was tested (*slope proportion*). This measure performed clearly worse than the *range proportion*, although yielding a significant difference between the two conditions (F=9.029, *df*=4/104, p=0.003). In the further analyses only the *range proportion* is included.

//////////INSERT TABLE 1 ABOUT HERE//////////

Within the two conversational conditions a number of syntactic configurations were tested.
In the turn-yielding sample, there is no significant effect due to syntactic structure and related aspects. By contrast, as regards the turn-holding condition, *range 1* and the *total range* differ significantly across the syntactic environments within the sample (F=4.276, *df*=3, p=0.008 for *range 1*; F=3.067, *df*=3, p=0.033 for the *total range*). A closer look at the data shows, however, that both effects are strongly influenced by a single instance, which has an exceptionally large *range 1* of 14.19 semitones. When this value is excluded, the statistics changes to F=2.728, *df* =3, p=0.051 for *range 1*, and F=2.233, *df*=3, p=0.092 for the *total range*. Thus, there are no clear effects due to the syntactic environment itself. Therefore, as in the turn-yielding condition, the sample is treated as independent of syntactic aspects.

3.2 Intercorrelations of the variables
In Table 2, the correlations are given for the eight acoustic features, combining the two conditions turn-yielding and turn-holding. The *total range* and the *range proportion* are statistically independent (r=--0.153). For the two *partial ranges* as well as for the two *partial durations* relatively high positive and negative correlations are found with the *range proportion* and the *duration proportion*, respectively. This follows from the definition of these indices. Yet, the partial ranges *range 1* and *range 2* are not correlated with each other --

i.e., in terms of regression, the shape of a rise cannot be predicted from its beginning. But, looking at the distribution of the values of *range 1* and *range 2* for the two conversational conditions, it becomes obvious that nonetheless there is some predictability. The scatterplot of Figure 4 shows systematic distribution differences for the two conditions: (a) in turn-yielding, there are no points for *range 1* > 6 st, nor for *range 2* < 2 st, but there are points for *range 2* > 10; (b) in turn-holding, there are no points for *range 2* > 10 st, but there are a substantial number of points for *range 2* < 2 st and for *range 1* > 6 st. This indicates that the separation of the two clusters of values can be captured by variables combining *ranges 1* and *2* (rather than by their correlation), e.g. by the range proportion (*range1*/[*range1* + *range2*]), which is supported by the inferential statistical analysis. It speaks in favour of treating the contours as unitary rather than composite. There were also significant correlations betweeen pitch-related and duration-related variables, e.g. *range proportion* and *duration proportion*. But these correlations are low.

/////////INSERT TABLE 2 ABOUT HERE//////////

/////////INSERT FIGURE 4 ABOUT HERE//////////

3.3 Contribution of the variables to the differentiation of melodic patterns for the two conditions (discriminant analysis I)

The contributions of the variables to the differentiation of the conversational conditions was analysed in two discriminant analyses. Additionally, the overall classification rate was determined for all variables together. At first, four features were analysed: *total range*, *range proportion*, *total duration*, and *duration proportion* -- despite the suppressor effects already indicated in connection with the formation of the *total range* by the combination of the partial *ranges 1* and *2*.

As shown in the left column of analysis I in Table 3, the discriminant analysis resulted in a significant discriminant function (Eigenvalue=0.321; CR=0.493; Wilk's lambda=0.757; *Chi-square*=28.388; *df*=4; p<0.0005). The *range proportion* makes the largest contribution to the differentiation of the two conversational conditions. The correlation with the corresponding canonical variable defined by the discriminant function is 0.992 (referred to as loading of the dependent variable on the discriminant factor). The *total range* contributes only slightly (with a loading of --0.237). For this variable the F-test was not significant. Standardized discriminant function coefficients and loadings of the *total duration* and the *duration proportion* are also small. Therefore, in this analysis, temporal factors are not important for the description of contours in the two conversational conditions.

/////////INSERT TABLE 3 ABOUT HERE//////////

/////////INSERT TABLE 4 ABOUT HERE//////////

The classification of cases achieved by the discriminant function of analysis I (cf. Table 4) produces the fewest errors with the patterns found for the turn-yielding condition (25% error rate). The prediction rate was worse for the turn-holding condition (31.1% error rate). The overall prediction was 70.8%. The clear division between the patterns from the two conversational conditions can be seen from the graph of the frequencies for the *range proportion* (Fig. 5a). The distribution shows a gradual descent to the right for the contours found within the turn-yielding condition. Only four cases exceed a value of 0.5. For the turn-holding condition the distribution is broader. There is a gradual ascent to the right with a number of cases exceeding values of 0.6. The poor performance of the *total range* can be seen from Figure 5b.

/////////INSERT FIGURE 5A AND B ABOUT HERE/////////

3.4 Alternative evaluation using ranges 1 and 2 (discriminant analysis II)
Because of the heterogeneity of the range feature, another discriminant analysis was calculated as a control. In this analysis, the *total range* was replaced by the two partial ranges (*range 1* and *range 2*). Similarly, the *total duration* was split up into *duration 1* and *duration 2*. The proportional variables were kept. It can be seen from the column of analysis II in Table 3 that again a significant discriminant function resulted (Eigenvalue=0.379; CR=0.524; Wilk's lambda=0.725; *Chi-square*=32.445; *df*=6; p<0.0005). As opposed to the *total range*, the partial ranges get considerable loadings (--0.586 for *range* 1 and 0.775 for *range 2*). But they do not reach those of the *range proportion* (--0.913). A further parameter with a relatively strong loading is *duration 1* (--0.495). As already shown, it is related to the *ranges 1* and *2,* although this relation is not close.

The inclusion of six variables in the analysis improved the prediction from 70.8% to 74.5% correct classifications based on the discriminant function. Misassignments decrease in the turn-holding condition. But in the new analysis the requirement of homogeneous variance-covariance matrices is again clearly missed (p<0.0005 for the Box-test). In this connection, speaker effects must be considered. In two pairs of male speakers a clear prediction of the communicative configurations was not possible -- because of idiosyncratic pitch patterns or characteristics of their speaking style. In the turn-holding condition, one speaker preferred contours with a large step between an almost flat low section on the accented syllable and a slightly rising one in the following stretch. Another speaker used high rising contours without a final gliding movement, one containing the extreme *range 1* value mentioned in 3.1. Two of these speakers have a South West German background, although there seems to be no direct dialect transfer. Excluding these speakers increases the rate of correct classifications to 80.5%; influences of durational factors disappear and the Box-test is no longer significant.

4 Discussion

4.1 Résumé of the findings
Our corpus data confirm that early valleys of KIM divide into two clearly different subtypes depending on the conversational condition they occur in. The difference between the two subtypes can best be caught by the feature of *range proportion*, which determines the location of the centre of a rising contour within a frame established by the *total range* of the contour. This feature behaves as predicted from the examples in Figure 1a, b. In the turn-yielding condition the *range proportion* is lowered (with an average range proportion of 0.35). In the turn-holding condition it is raised to an average range proportion of 0.54. Being determined by the interrelation of two stretches within the contour, the one containing *range 1* and the other *range 2*, the *range proportion* is a property of the whole accent-related pattern. This pattern begins before the accented-vowel onset, which is bridged by *range 1*.

The "proportional property" of a rise does not imply that from the first part of the rising movement a second part can immediately be calculated (i.e., *range 2* from *range 1*). Rather, the range of the first part of the rise (*range 1*) in relation to the *total range* must not exceed, or fall below, a critical value, to get an unambiguous assignment to communicative function. Within this limit, *range 2* can take different values. This particular behaviour of the variation of *ranges 1* and *2* may be the reason why their correlation is not negative (although this could be expected on the basis of the means of *ranges 1* and *2* for the two conversational conditions).

Combining the means of *ranges 1* and *2*, average contours can be drawn for both the turn-yielding and the turn-holding conditions (Fig. 6). These contours illustrate the effect of a different *range proportion* on the course of a contour -- also accounting for the *duration proportion* found. An average contour curved inwards results for the turn-yielding condition, a contour curved outwards results for the turn-holding/continuing condition. The *slope proportion*, as an alternative measurement of the contour shape is clearly less suited to describe differences between the conditions since it is affected by syllabic or segmental complexity.

/////////INSERT FIGURE 6 ABOUT HERE/////////

A further variable characterizing rises as whole patterns is the *total range*, although this variable gives no direct account of the shape of rising contours. The *total range* is often taken to be the decisive feature distinguishing rises in question vs. continuation contexts. But in our study it does not contribute much to the distinction of the conditions compared. The relatively small contribution of the *total range* can again be gathered from Figure 6 showing a mean difference of only 0.9 semitones. However, excluding four speakers this mean difference increases to 1.7 semitones (cf. 3.1). Thus, there are indications that the contribution of the *total range* may be underestimated in our data.

Unlike the *total range*, *range 1* and *range 2* are important variables in our data set getting high loadings and improving the classification in the second discriminant analysis. This is particularly true of *range 2*. For this variable, the loading comes close to loadings found for the *range proportion*, and the mean differences between the two conditions amount to 2.5 semitones (3.1 semitones without four speakers). Both results suggest that there may be a separate tonal or melodic unit governing the pitch course at the end of a prosodic phrase, e.g., a particular boundary tone. On the other hand, although the relation between the partial ranges cannot be caught by a statistical correlation, increases of *range 2* benefit from decreases in *range 1*. That is, even if we were to accept that the pattern observed in the turn-yielding condition as opposed to the one in the turn-holding condition is marked by a certain tonal event concluding the contour and causing high-rising phrase-final pitch, this feature remains a property of the contour as a whole.

Thus, *the total range* of a rise seems to be a compensatory system of interacting stretches. Although the range as a whole contributes only slightly or not at all to the functional distinction, it provides the frame for the variation of the shape of rises. For early valleys the rising movement included in this frame already begins before the onset of the accented vowel (cf. Fig. 3). The turning point of the contour or the limits of the interacting stretches are only provisionally determined in our analysis (cf. 2.1).

Finally, two issues are addressed that may affect the validity of the conclusion to be drawn. The first concerns the role of *duration*. There is a small, but significant correlation between range proportion and duration proportion – although these variables should be independent. In particular, as a between-groups effect, the larger step towards the pivot in the turn-holding condition also takes more time. Effects of this kind may be influenced by the (provisional) positioning of the turning point at the F0 maximum of the accented syllable of polysyllabic accent groups. But variation due to the position of the turning point can only be expected *within* the conversational conditions (see below), because the conditions are quite homogeneous with respect to syllabic structure. However, by and large the covariation between pitch and duration is small compared with the pure pitch effects found.

The second issue concerns the *pattern variation within the conversational conditions* of turn-holding and turn-yielding. Not accounting for this within-variation can obscure the character of the variation between the conditions. Five sources of variance may be of importance: (1) variance due to segmental and syllabic complexity (and to the definition of the pivot point), (2) variation due to syntax (here no significant variation within the conditions was found), (3) variation due to speaker characteristics, (4) dialectal variation, (5) variation due to communicative function, assuming that turn-yielding vs. turn-holding only captures part of an underlying concept or shows a probabilistic assignment to this concept. Above all, the turn-holding condition reveals some pattern variation. In spite of these factors contributing to error variances, the results of the analysis are quite clear. The question of communicative inhomogeneity will be discussed in section 4.3.

4.2 Conclusions with respect to form

The rising patterns examined in the present study are cases of a uniform contour type: they are early valleys in terms of KIM. But allocating early valley contours to two different conversational conditions results in two different *sub-patterns*. Acoustically, these sub-patterns are above all characterized by a particular contour feature: the *range proportion*, giving a better characterization of the types of rising movements than does the *total range*, or the second part of the range (*range 2*), or measures based on slopes. Different from the mere excursion or height of phrase-final rising, the *range proportion* is a marker of integral melodic patterns. The results point to two such patterns which may be viewed as communicative gestures used in discourse. Being integral also means that these patterns can only insufficiently be accounted for as sequences of single tonal events. Rather, they are coherent gestalt-like shapes in which all elements are related and can only be determined by their connection.

The frequency distributions of the measurements found for the *range proportion* in the two conversational conditions show an overlap, with a decrease in the one distribution being paralleled by an increase in the other. This looks like a continuous transition between the observed patterns. However, in trying to assess the character of the pattern variation between the conditions it should be taken into account that there is a number of factors causing variance within them and perhaps obscuring effects. On the whole, the profile of the data is so clear that we may expect separate *prototypical patterns* as an underlying structure.

The results reported here are closely linked to the underlying approach that views accent-related pitch courses as integral contour units related to the accented-vowel onset. This approach enabled us to see the two contour shapes which are dealt with here in a *common perspective* – which, in turn, revealed their difference. Alternatively, from the autosegmental-metrical point of view, e.g. GToBI [cf. Grice and Baumann, 2000], it would have been much harder to work out this joint perspective. With the entire accented syllable as a reference for placing pitch accents, patterns of the kind illustrated by example (a) above (concave patterns) may get an L* on the accented syllable on the basis of a trough in the contour. Others would get an H*. For patterns such as in example (b) (convex patterns) H* may be preferred. Thus, the early valley in terms of the Kiel Intonation Model cannot be clearly distinguished in this framework. In consequence, convex and concave shapes in continuous rises would not be dealt with as different types of the same contour category. Further, in a tone sequence perspective, the end of rising contours would be treated independently by using phrasal and boundary tones. Thus, there are two competing ways to model the variation within the contour category analysed here, via pitch accents and via phrasal and boundary tones. This kind of a compositional instead of an integrating view can also be seen in the Dutch IPO perspective of rises [‘t Hart, Collier, and Cohen, 1990; cf. Caspers 1998]. Both approaches fail to recognize that the phrase-final gesture (with all its semantic colouring) already starts with the preceding accent-related pitch movement.

4.3 Conclusions with respect to function

The present analysis started with the distinction between two conversational conditions: turn-yielding and turn-holding, which prove to be quite suited to separate pitch patterns of the kind portrayed in the examples in Figure 1. The profiles sketched with the help of the average ranges and durations (cf. Fig. 6) point to these prototype-like contours. This suggests that patterns differing in this way may be used in conversation as means to signal discoursal functions relevant to turn-yielding and turn-holding.

However, in both conditions there are several cases that clearly deviate from the contrasting patterns. Consequently, the empirical profiles are less explicit than expected. The variation observed in each of the two conversational conditions decreases the rate of correct classifications calculated from the discriminant functions. Potential sources of variation in the conditions have already been mentioned above. Among these, effects of communicative inhomogeneity are of particular interest, as they may influence the choice or explicitness of melodic patterns. Inhomogeneity means that turn-yielding and turn-holding may only give an incomplete account of the communicative meaning conveyed by the two prototypical contour shapes. The following arguments are to lead to a more accurate definition of the functional concept differentiating phrase-final early valley contours.

The formal and phonetic variation of rising contours may be discussed in connection with *question* and *continuation* [e.g. Caspers, 1998, 2001; Gilles, 2000]. However, question and continuation are only partly suited for the functional classification of the pattern variation discussed here. On the one hand, these terms are not evenly matched. They refer to concepts which do not share the same functional dimensions. Questions are frequently defined via syntax; the counterpart is statement. Continuation is defined negatively, via incompleteness, which may be syntactic or communicative [cf. Auer, 1996]; the counterpart is conclusion. On the other hand, in German, both pitch patterns considered here can occur with question and continuation, and also with other types of utterances [Gilles, 2000] -- although the frequencies are different. Moreover, other rising, as well as falling, patterns can be used in questions [Essen, 1964; Kohler, 2004; Selting, 1995] and continuation [Meinhold, 1967; Essen, 1964; Gilles, 2000, 2003]. The same applies to utterances in turn-yielding and turn-holding contexts [Selting, 1995; Gilles, 2003].

Yet, a more general and at the same time more specific notion of the function of different shapes in rising contours can be derived from a *synopsis of functional terms* that are used in work on question and continuation and on rising intonation for German. This synopsis reveals two polar syndromes of semantic features that can be taken as a basis for a communicative background concept controlling the use of concave vs. convex patterns:

establishing contact [Stock and Zacharias, 1971; Mixdorf, 2004; Meinhold and Stock, 1980], *showing openness directed towards the other communicator* [Kohler, 1991a; Féry, 1993], *conveying concern for the co-participant* (Kohler 1991a), *leaving alternatives* [Kohler, 1991a; Essen, 1964], *turn-yielding* [Selting, 1995]
vs.

continue speaking [Gilles, 2003], *direction towards the speaker* [Féry, 1993], *turn-holding* [Selting, 1995; Auer, 1996].

The applicability of these meaning features is supported by the communicative interpretations of the examples in this paper and in Dombrowski and Niebuhr [2005]. As a generic functional dimension encompassing the various meaning terms, we propose *activation* vs. *restriction*, which denotes two basic communicative sets or activities controlling interpersonal interaction. A person *activating* another person tries to convey that the recipient has alternatives (i.e., behavioural or conversational alternatives). In turn, *restricting* means to convey that behavioral or conversational choices are reduced. The first may be the case when a speaker tries to establish contact, or gives another speaker freedom for activities, e.g. leaving more than one alternative for answering a question, or yielding the turn. The second may be the case if the other speaker is simply to listen and must be prevented from taking the turn, or if s/he is required to react in a particular way. This communicative dimension may help to explain why sometimes in turn-yielding or in questions with rising intonation valley contours are convex, and perhaps also why they may sometimes be concave in turn-holding or continuation. In Dombrowski and Niebuhr [2005] examples are given that provide casuistic evidence in favour of this view.

4.4 Gestural character of the patterns analysed

Both patterns considered in the above corpus analysis are characterized by a rising course. Thus they make use of the semantics of high pitch. On the one hand, high pitch can express tension [as demonstrated in studies by Scherer and co-workers, cf. Scherer, Johnstone, and Klasmeyer, 2003; cf. also Gussenhoven's effort code: Gussenhoven, 2004]. On the other hand it can express dependence on other communicators or consideration or high regard which can all be related to Ohala's [1983, 1984] frequency code. In examining the communicative background of turn-yielding vs. turn-holding and question vs. continuation, a further aspect in the relation of pitch and meaning emerges, which is, however, not independent of the two aspects just mentioned. We have pointed out two patterns of *complex F0 time shapes* with two different functions, characterized by *activation* vs. *restriction*. One pattern raises pitch on the accented syllable and at the same time restricts the further rise after the accented-vowel onset. The other pattern lowers pitch on the accented syllable (within the global rising movement) giving an enlarged frame for the subsequent part of the rise.

Within the functional frame outlined in section 4.3, the patterns analysed in this paper may be regarded as two *communicative gestures*. The first one can be understood as a signal activating the other participants in a discourse. The German term *Kontaktintonem* [contact intoneme; Stock and Zacharias, 1971; Meinhold and Stock, 1980] reflects this communicative gesture. Signals of this type can occur in various conversational configurations, they are not restricted to the utterance type of question. The gliding movement often found with a low range proportion may be related to universal aspects of communication, going beyond the high-low distinction implied in the frequency code. The particular quality of the rising movement would have to be added.

The second pattern is a gesture of restriction, imposing a restraint on a listener who is already an attentive co-participant in a discourse. This "external" restricting function at the same time structures a speaker's turn. Therefore the term "self-directed" may be justified. The restricting gesture is often associated with the linguistic function of continuation, but it can occur in other communicative configurations as well [Dombrowski and Niebuhr, 2005].

So, a concept of *activation* vs. *restriction* may be useful for a clearer description of interpersonal interaction in conversation analysis and function-based work on prosody.

Acknowledgements
We would like to thank Klaus Kohler for generously giving us his time and advice in the course of our work. The discussions with him contributed a great deal to improving the paper. We also thank the two reviewers John Kingston and Francis Nolan for very helpful comments and criticisms.

References

Adriaens, L.M.H: Ein Modell deutscher Intonation; doct. diss. TU Eindhoven (1991).
Auer, P: On the prosody and syntax of turn-continuations. In Couper-Kuhlen; Selting, Prosody in conversation: Interactional Studies, pp. 57--100 (Cambridge University Press, Cambridge 1996).
Brosius, F: SPSS 11. (mitp, Bonn 2002).
Caspers, J.: Who's next? The melodic marking of question versus continuation in Dutch. Language and Speech *41*: 375--398 (1998).
Caspers, J.: Testing the perceptual relevance of syntactic completion and melodic configuration for turn-taking in Dutch. Proc. Eurospeech 2001, Aalborg, Denmark (2001).
Dombrowski, E; Niebuhr, O.: Phrase-final rises in German: some examples. URL http://www.ipds.uni-kiel.de/kjk/publikationen/audiobsp.en.html. (2005).
Essen, O. von: Grundzüge der hochdeutschen Satzintonation (A. Henn, Ratingen 1964).
Féry, C.: German Intonational Patterns (Max Niemeyer, Tübingen 1993).
Fowler, C.A.;Brown, J.M.: Intrinsic f0 differences in spoken and sung vowels and their perception by listeners. Perception and Psychophysics *59*: 729--738 (1997).
Gartenberg, R.;Panzlaff-Reuter, C.: Production and perception of f0 peak patterns in German. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel *25*: 29--115 (1991)
Gilles, P.: Intonation der Weiterweisung. InList *20*: 1--35 (2000).
Gilles, P.: Die Intonation von Abschluss und Weiterweisung in deutschen Regionalvarietäten: Funktionale und phonetische Analysen; habil. diss. University of Freiburg (2003).
Grabe, E.: Pitch accent realization in English and German. Journal of Phonetics *26*: 129--144 (1998).
Grice, M.;Baumann, S.: Deutsche Intonation und GToBI. Linguistische Berichte *191*: 267--298 (2000).
Gussenhoven, C.: The phonology of tone and intonation (Cambridge University Press, Cambridge, UK 2004).
‘t Hart, J.; Collier, R.; Cohen, A.: A perceptual study of intonation: An experimental-phonetic approach to speech melody. (Cambridge/New York, Cambridge University Press 1990).
IPDS: The Kiel Corpus of Read Speech, Volume 1, CD-ROM#1. (Institut für Phonetik und digitale Sprachverarbeitung, Kiel 1994).
IPDS: The Kiel Corpus of Spontaneous Speech, Volume 1, CD-ROM#2. (Institut für Phonetik und digitale Sprachverarbeitung, Kiel 1995).
IPDS: The Kiel Corpus of Spontaneous Speech, Volume 2, CD-ROM#3. (Institut für Phonetik und digitale Sprachverarbeitung, Kiel 1996).
IPDS: The Kiel Corpus of Spontaneous Speech, Volume 3, CD-ROM#4. (Institut für Phonetik und digitale Sprachverarbeitung, Kiel 1997).
Kohler, K.J.: Categorical pitch perception. Proc. 11th ICPhS 1987, Volume 5, Talinn, Estonia: 331--333 (1987).
Kohler, K.J.: A model of German intonation. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung *25*: 295--360 (1991a).

Kohler, K.J.: Prosody in speech synthesis: the interplay between basic research and TTS application, Journal of Phonetics *19*: 121--138 (1991b).
Kohler, K.J.: Terminal intonation patterns in single accent utterances of German: Phonetics, Phonology, and Semantics. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung *25*: 115--186 (1991c).
Kohler, K.J.: PROLAB – the KIEL system of prosodic labelling. Proc. 13th ICPhS 1995, Volume 3, Stockholm, Sweden: 162--165 (1995).
Kohler, K.J.: Modelling prosody in spontaneous speech. In Sagisaka; Campbell; Higuchi, Computing Prosody: Computational Models for Processing Spontaneous Speech, pp. 187-210 (Springer, New York/Berlin 1997).
Kohler, K.J.: Pragmatic and attitudinal meanings of pitch patterns in German syntactically marked questions. International Symposium on Tonal Aspects of Languages: Emphasis on Tone Languages (TAL), Beijing (2004).
Kohler, K.J.; Pätzold, M.; Simpson, A.P.: From scenario to segment: The controlled elicitation, transcription, segmentation and labelling of spontaneous speech, Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel *29*: 1--137 (1995).
Landgraf, K.: Steigende Intonationskonturen im Deutschen – Experimentalphonetische Untersuchungen zur auditiven Kategorisierung; M.A. diss. University of Kiel (2003).
Lehiste, I.: Suprasegmentals (M.I.T. Press, Massachussetts 1970).
Meinhold, G.: Progrediente und terminale Intonationsverläufe im Deutschen. ZPSK *20*: 465--478 (1967).
Meinhold, G.; Stock, E.: Phonologie der deutschen Gegenwartssprache. (Bibliographisches Institut, Leipzig 1980).
Mixdorff, H.: Quantitative analysis of prosody in task-oriented dialogs. Proc. Speech Prosody 2004, Nara, Japan: 283--286 (2004).
Niebuhr, O.: Perceptual study of time variables in F0 peaks. Proc. 15th ICPhS 2003, Barcelona, Spain: 1225--1228 (2003).
Niebuhr, O.: Intrisic pitch in opening and closing diphthongs of German. Proc. Speech Prosody, Nara, Japan: 733--736 (2004).
Niebuhr, O.; Kohler, K.J.: Perception and cognitive processing of tonal alignment in German. Proc. International Symposium on Tonal Aspects of Languages: Emphasis on Tone Languages (TAL), Beijing, China: 155--158 (2004).
Ohala, J.J.: Cross-language use of pitch: An ethological view. Phonetica *40*: 1--18 (1983).
Ohala, J.J.: An ethological perspective on common cross-language utilization of f0 of voice. Phonetica *41*: 1--16 (1984).
Peters, B.; Kohler, K.J.: Trainingsmaterialien zur prosodischen Etikettierung mit dem Kieler Intonationsmodell KIM. Unpublished manuscript, IPDS, University of Kiel (2004). URL http://www.ipds.uni-kiel.de/pub_exx/bpkk2004_1/trainerA4.pdf
Scherer, K.R.; Johnstone, T.; Klasmeyer, G: Vocal expression of emotion. In Davidson; Scherer; Goldsmith, Handbook of Affective Sciences, pp. 433-480 (Oxford University Press, Oxford 2003).
Selting, M.: Prosodie im Gespräch (Max Niemeyer, Tübingen 1995)
Stock, E.: Untersuchungen zu Form, Bedeutung und Funktion der Intonation im Deutschen (Akademie, Berlin 1980).
Stock, E.; Zacharias, C.: Deutsche Satzintonation (Verlag Enzyklopädie, Leipzig 1971).
Tatsuoka, M.M.: Multivariate Analysis (Wiley, New York 1971).

Table 1. Means and standard deviations, as well as F-ratios, degrees of freedom, and probabilities for the tests of mean differences for each of the eight acoustic variables measured in the phrase-final early valleys.

	turn-yielding		turn-holding		F	Df1 / df2	p
	mean	s	Mean	s			
range 1 (st)	3.43	1.30	5.04	2.32	13.520	1 / 104	<0.0005
range 2 (st)	6.77	2.59	4.29	2.33	23.674	1 / 104	<0.0005
total range (st)	10.20	2.47	9.33	3.21	1.867	1 / 104	0.175
range proportion	0.35	0.13	0.54	0.17	32.812	1 / 104	<0.0005
duration 1 (ms)	235	85	311	126	9.643	1 / 104	0.002
duration 2 (ms)	302	199	298	139	0.011	1 / 104	0.971
total duration (ms)	537	230	609	188	2,879	1 / 104	0.093
duration proportion	0.46	0.14	0.51	0.14	2.861	1 / 104	0.094

Table 2. Correlation matrices for all acoustic features measured. PM correlations and significance levels.

n=106	*range 1*	*range 2*	*total range*	*range proportion*	*duration 1*	*duration 2*	*total duration*	*duration proportion*
range 1	1	-0.102	0.605***	0.654***	0.264**	0.131	0.255**	0.005
range 2		1	0.730***	-0.753***	-0.305**	0,029	-0.152	-0.310
total range			1	-0.153	-0.063	0.114	0.054	-0.245**
range proportion				1	0.367***	0.044	0.245**	0.220*
duration 1					1	0.046	0.610***	0.623***
duration 2						1	0.820***	-0.691***
total duration							1	-0.191*
duration proportion								1

Table 3. Discriminant analyses I and II. Coefficients of the discriminant functions and centroids for the functions, correlations between the dependent variables (acoustic features) and the canonical variables defined by the discriminant function (loadings). Probabilities and significances for the Box tests and for the discriminant functions.

	analysis I coefficients	analysis I loadings	analysis II coefficients	analysis II loadings
range1 (st)			-0.243	-0.586
range2 (st)			0.618	0.775
total range (st)	-0.077	-0.237		
range proportion	0.941	0.992	-0.247	-0.913
duration1 (ms)			-0.902	-0.495
duration2 (ms)			0.872	0.017
total duration (ms)	0.083	0.294		
duration proportion	0.083	0.293	1.142	-0.269
centroids turn-yielding:	-0.853		0.927	
centroids: turn-holding	0.369		-0.401	
Box test (p)		0.012*		0.000***
discriminant functions (p)		0.000***		0.000***

Table 4. Classification achieved by discriminant analysis I and II, the latter in parentheses. The percentage of correct classifications is 70.8% in analysis I and 74.5% in analysis II.

			predicted group: turn-yielding condition	predicted group: turn-holding condition	total
original group	frequencies	turn-yielding condition	24 (24)	8 (8)	32
original group	frequencies	turn-holding condition	23 (19)	51 (55)	74
original group	percentages	turn-yielding condition	75.0 (75.0)	25.0 (25.0)	100,0
original group	percentages	turn-holding condition	31.1 (25.7)	68.9 (74.3)	100,0

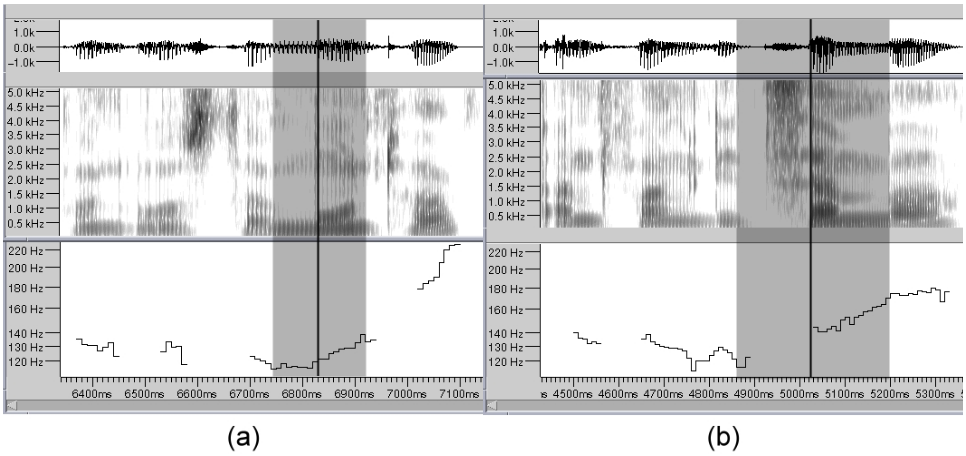

Figure 1. Oscillogram, spectrogram and F0 course of the sections *und Osterm'ontag* and *Anfang Dez'ember* from example (a) and example (b). Accented syllables are highlighted with grey, the vertical lines indicate the accented-vowel onsets. As suggested by the F0 course and confirmed by listening, the rising pitch movement already starts before the accented-vowel onset in both cases. But the rise takes different shapes, with a higher F0 level reached at the end of the accented syllable in example (b) as opposed to example (a).

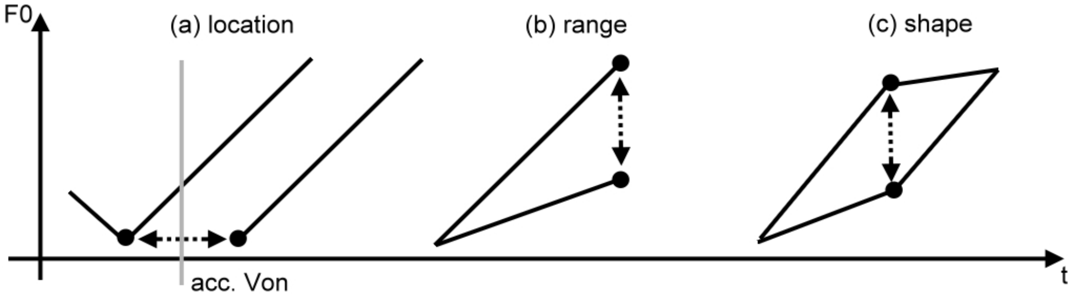

Figure 2. Three dimensions describing phrase-final rises applying to early and late valleys of KIM: location (a), range (b) and shape (c).

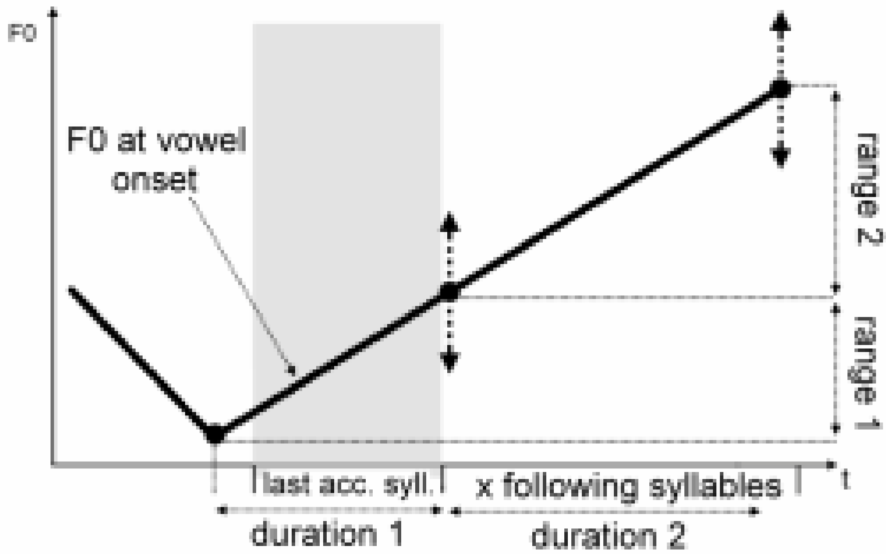

Figure 3. Contour model of the phrase-final early valley.

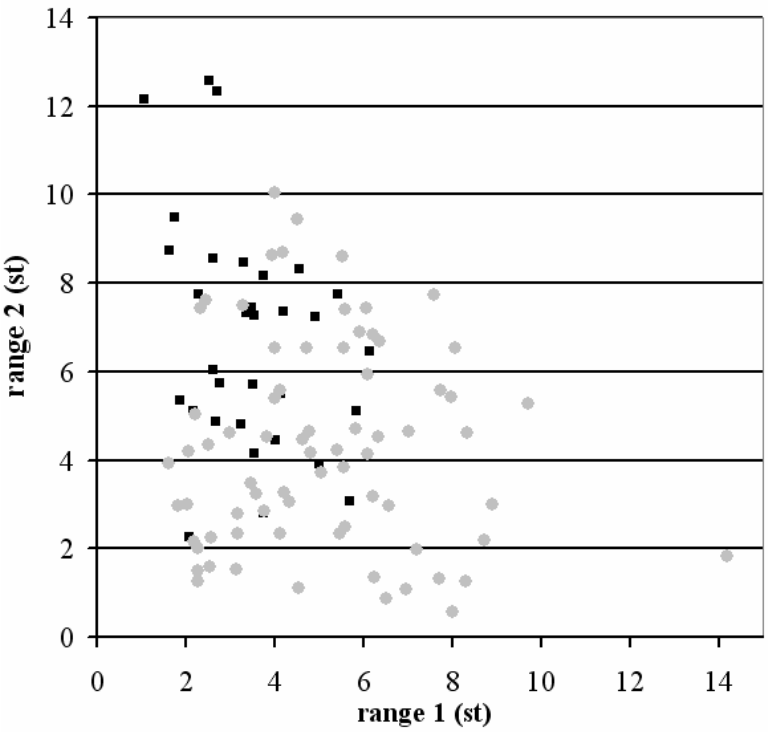

Figure 4. Scatter plot showing the type of covariation found for *range 1* and *range 2* in early valleys. Black squares indicate the turn-yielding and grey circles the turn-holding condition. Values are given in semitones (st).

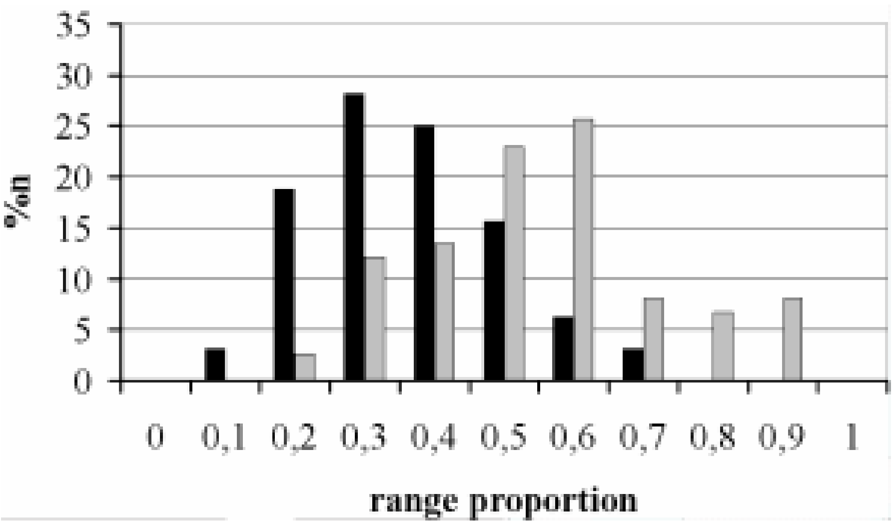

Figure 5a. Relative frequencies of the values for the *range proportion* in the turn-yielding (black) and the turn-holding (grey) conditions. The *range proportion* is scaled in tenths ranging from 0 to 1.

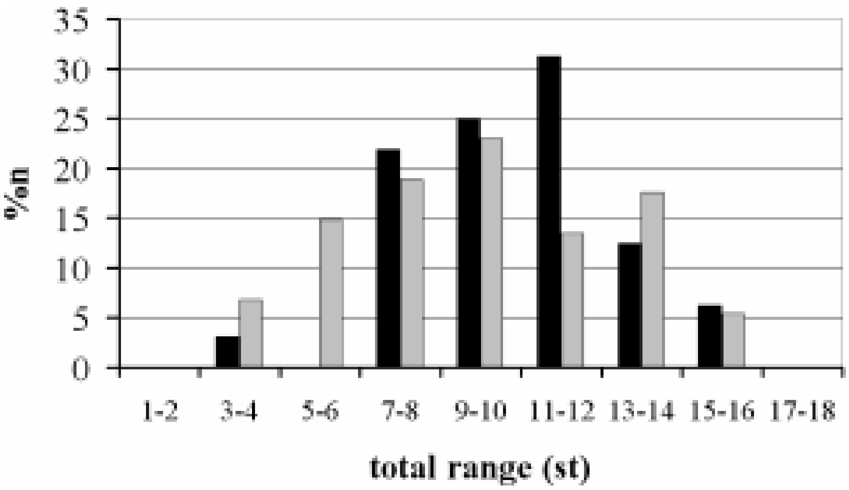

Figure 5b. Relative frequencies of the values for the *total range* in the turn-yielding (black) and the turn-holding (grey) conditions. The *total range* is scaled in steps of two semitones.

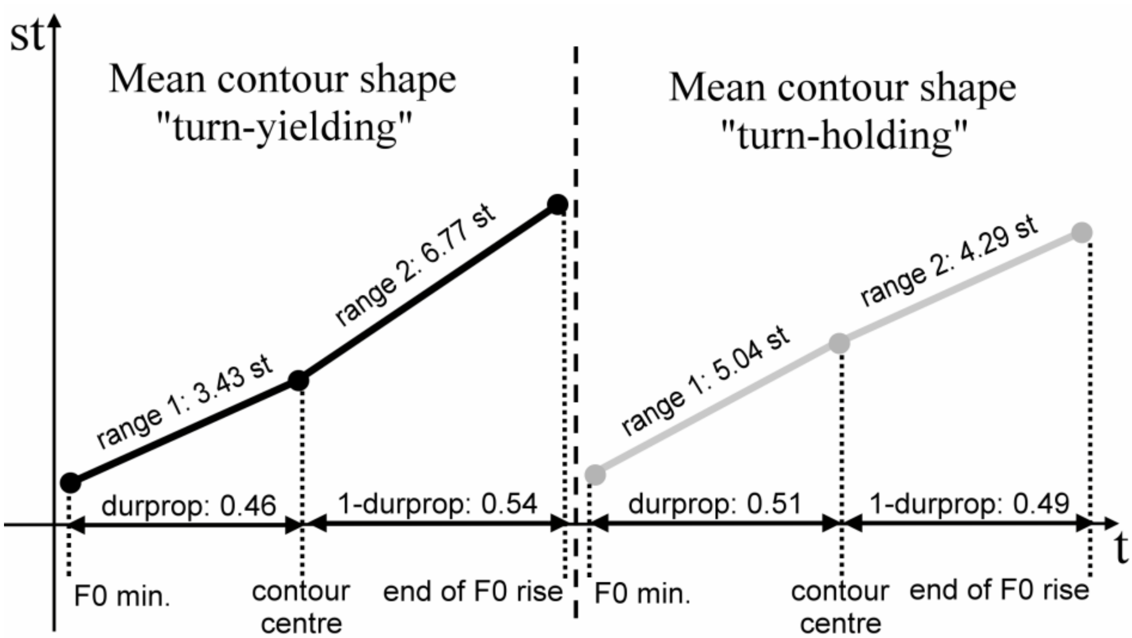

Figure 6. Mean contour shapes of early valleys in the turn-yielding and the turn-holding conditions, according to the values of Table 1. Since the two semitone distances in each contour (*range 1* and *range 2*) have almost equal mean durations (i.e., *duration proportion*s of 0.46 and 0.51), the resulting contour shapes are characterized by a concave or convex course, respectively. The starting pitch of both mean contour shapes is variable. Only their contour is shown, its position within the speaker's average range of voice has not been accounted for.