Klaus J. Kohler (Kiel)

Paradigms in Experimental Prosodic Analysis: From Measurement to Function

1 Introduction

Theoretical and methodological paradigms in speech research determine the design of data collection, their analysis and their interpretation. The scientific approaches to spoken language have been shaped by the dominating influence of the dichotomy of phonology, dealing with discrete mental objects, and phonetics, dealing with infinitely fine gradation of physical manifestation. They have been associated with the humanities and the sciences, respectively, since the first half of the 20th century. It is still wide-spread among linguists to look at phonology as a level of purely symbolic representation without reference to the physics of speech, and on the other hand, for engineers to apply signal processing techniques to human speech without reference to its linguistic and communicative functions. These two camps are so diametrically opposed to each other that understanding between them is impossible and their separate contributions to the elucidation of speech communication are of dubious value.

But an increasing number of phonologists take their phonological solutions to the laboratory to fill them post hoc with phonetic measurement. This has become the new paradigm of Laboratory Phonology, as represented by the conference and proceedings series with that name. Under the sociology of science perspective, it constitutes today’s dominant paradigm. In this rapprochement of the two opposed fields of study, linguistic form takes precedence over phonetic substance and is the structural basis for, and unaffected by, the latter. Furthermore, function is also subordinated to form and restricted to linguistic function, thus creating the dichotomy linguistic vs paralinguistic. So, linguistic form assumes a pivotal mediating role between substance and function, excluding direct links.

This is the theoretical background for phonetic measurement in highly stylized language material. Segmental analysis is based on words (even logatomes) in isolation or in metalinguistic sentence frames, and prosodic analysis

*The development of the Kiel Intonation Model presented in this paper has benefited from discussions with Oliver Niebuhr and Ernst Dombrowski in the author’s Prosody Colloquium at IPDS Kiel. Oliver Niebuhr also helped with graphic formatting. Their contributions as well as helpful suggestions from two anonymous reviewers are gratefully acknowledged.
on isolated, non-contextualized sentences, often of doubtful semantics. The
generalizations from such laboratory data to phonetic patternings in speech
communication are quite limited, which is nowhere more obvious than in the
field of prosody. Two examples are to illustrate the disregard for function,
contextualization and semantic as well as pragmatic plausibility in prosodic
data collection within the framework of Laboratory Phonology.

(1)   Die Nonne und der Lehrer wollen der Lola in Murnau eine Warnung geben, und die Hanne
       will im November ein Lama malen.
       ’The nun and the teacher want to give a warning to Lola from Murnau, and Hanna wants to
       paint a lama in November.’ (Truckenbrodt, 2002)

(2)   (a)   Malen die Schüler die Wohnungen?
         ’Are the pupils painting the flats?’
   (b)   Hatten die Schüler die Wohnungen bemalen wollen?
         ’Had the pupils wanted to decorate the flats?’
       (Grice et al., 2000) The translations are the authors’ and reflect the assumption that both the
       noun ‘Maler’ and the verb ’malen’ refer to both ‘decorator’ and ‘artist’. But German usage
       does not connect the verb ’malen’ with ‘decorating’, and ’bemalen’ conjures up a picture of
       vandals breaking into flats and spraying the walls with their spray cans.

In these cases, sentences are constructed in such a way that the relevant parts
under investigation only contain sonorous sounds in order to avoid micro-
prosodic interference with the macroprosodic intonation contours. But the
regularities of this interaction are quite well understood today, and intonation
contrasts should therefore be studied expressly in different segmental settings,
as was done by Gartenberg and Panzlaff-Reuter (1991). This allows using
more natural sentences for the analysis of, e.g., upstep and phrase accents.
These topics cannot be investigated with the above type of artificially con-
structed material if the results of measuring F0 curves in production data are to
throw light on how these prosodic features operate in natural interaction,
which prosodic analysis needs to achieve.

In spite of phonetic measurement in phonological categories, Laboratory
Phonology perpetuates the old phonology-phonetics dichotomy because it
continues to ask the question whether the difference between two utterances IS
phonological, i.e. belongs to “discrete cognitive representations”, or IS pho-
netic, i.e. is “variation in their physical implementation” (cf. Gussenhoven, this
volume). This de Saussurian question was aggravated by the very innovation
of Laboratory Phonology, viz. the inclusion of gradient measurement in dis-
crete categories. So it becomes a vexing task to establish discreteness in grad-
ient variability, especially in the field of intonation. Pitch variability is said to
be paralinguistically meaningful, but this type of semantics is at the same time
seen as hiding the fundamental cognitive structures, the ultimate goal of lin-
guistic analysis. Pitch accents and boundary features are represented as sequences of L and H tones, which are regarded as existentially given discrete categories. Phonetic variability needs to be projected onto them without residue. On the other hand, H maximum and L minimum, and F0 alignment are seen as gradually scaled and consequently not linguistic. If they add to the meaning of utterances it must be paralinguistic.

This primacy of linguistic form over substance and function has been questioned in the research by Björn Lindblom and John Ohala over many years (e.g. Lindblom, 1990; Ohala and Jaeger, 1986). Their paradigm may be called Function-Oriented Experimental Phonetics. In this paper, I shall be presenting this alternative paradigm to Laboratory Phonology with special reference to the analysis of prosody, on the basis of the Kiel Intonation Model (KIM). In its outline I make contrastive references to the Autosegmental-Metrical (AM) approach to intonation and its labelling framework ToBI within Laboratory Phonology (Pierrehumbert et al., 2000), whose determining features are:

(3) (a) the representation of global pitch patterns by local phonological tones (H or L), associated (in an abstract, atemporal way) with specific syllables and boundaries
(b) the post hoc projection of these tones onto the time scale through measurement of phonetic alignment of F0 maxima/minima with articulation
(c) the concentration on speech production and the neglect of the question as to the perceptual relevance of, e.g., alignment differences in acoustic signals within and across languages
(d) lack of verbal and situational context in data acquisition, analysis and interpretation.

2 The Kiel Intonation Model (KIM)

2.1 Tone contours vs tone levels

KIM (Kohler, 1991a,b, 1995, 1997) and its labelling tool PROLAB (Kohler, 1995, 1997; Peters and Kohler, 2004) are in the European tradition of intonation research, represented among others by the London School of Phonetics (e.g. O'Connor and Arnold, 1961), and are rooted in a system of global pitch contours (peak, valley, combined peak-valley and level patterns) and their syntagmatic concatenations. This contrasts with the American tradition of tone-level analysis, which was initiated by Pike (1945). The AM model and its labelling tool ToBI reduced Pike's 4-level system to two distinctive levels – high H and low L – and allocated them to three categories – pitch accents, phrase accents and boundary tones.

A rising-falling-rising pitch pattern in a non-committal, implicational “Yes.” in English, or in a friendly warning “Du!” in German (e.g. spoken to a child and accompanied by a movement of the raised index finger) is a global intona-
tion unit in contrast with a rising-falling or just a rising one, in categorical statements "Yes."/"Du." or questions "Yes?"/"Du?", respectively. These intonation units are produced and perceived holistically and associated as such with different semantics and pragmatics. KIM/PROLAB treat them as unitary elements in intonational phonology, viz. a combined peak-valley contour vs a peak contour vs a valley contour. AM/ToBI, on the other hand, need to split them up into sequences of three tone-level units: a pitch accent + a phrase accent of an intermediate phrase + a boundary tone of an intonation phrase.

The rising-falling-rising pattern requires four significant F0 points for its unique definition – the low start, the maximum, the subsequent minimum and the following higher end point. Traditional AM/ToBI allow the combination of only two tone levels in a pitch accent. (This restriction has never been justified theoretically; and the further limitation of bitonal possibilities was made plausible with reference to empirical data, thus creating an explanatory circle.) This means that "Du!" needs another two tonal points, which are supplied by the phrase and boundary tones. Whereas rising-falling "Du." and just rising "Du?" could be uniquely defined by one phrase tone in addition to the pitch accent tone(s), rising-falling-rising "Du!" needs a second phrase tone, which then has to be of a categorically different type, i.e. the phrase accent associated with a non-final intonation phrase. This leads to the general postulate of a sequence of a non-final and a final phrase tone for all final intonation phrases, viz. L+H*L-H%, L+H*L-L%, L*H-H% in "Du!", "Du.", "Du?", respectively. What resulted from the empirical constraint of descriptive adequacy in "Du!", became a systemic constraint in the other cases.

The unitary nature of the rise-fall-rise pattern in the discussed utterances is further supported by the fact that it can also occur phrase-externally. For example, the German sentence "War er das wirklich!?" ("Was it him really!?") may be pronounced with two pitch accents, one on "war" and another on "wirklich", a rise-fall-rise pattern across "War er das", and a separate rising pattern on "wirklich". The latter starts below the end point of the preceding combined contour, without a perceivable signal for an intervening syntagmatic prosodic phrase boundary (e.g. lengthening of "das"). This melodic sequence of fall-rise - fall + rise is perceptually and semantically differentiated from another two-accent pattern, namely (rising-)falling on "War er das", followed by rising on "wirklich". The rise-fall-rise expresses disbelief, whereas the (rise-)fall simply contrasts the polarity of the verb: "was it him or not?" There is also a perceptual and semantic difference between the absence and the presence of an acoustically signalled intervening phrase boundary between a rise-fall-rise in "War er das" and a rise in "wirklich": with such a boundary, "wirklich" becomes a second question. Finally, all three patterns are perceptually different from a one-accent rendering of the sentence, where the rise-fall-rise extends over the whole word sequence, starting on the accented "war". Unaccented "wirklich" loses its reinforced meaning "in reality" and becomes a modal par-
ticle. (Observations by trained phoneticians in the author’s Prosody Collo-
quium at IPDS Kiel.)

In the one-accent case, AM/ToBI would associate a pitch accent L+H* with
“war”, and L-H% with the intermediate and intonation phrase boundaries. In
the two-accent case of (rise-)fall + rise, a pitch accent L+H* would be associ-
ated with “war”, another pitch accent L* with “wirklich”, and again L-H% with
the boundaries. In the case of rise-fall-rise + phrase boundary + rise, the
first intonation phrase would contain L+H*+L-H%, the second L*H-H%. In the
case of no intervening phrase boundary, AM/ToBI are unable to represent the
empirically found pattern within the existing framework of tone-level se-
quen
ces. They would associate one pitch accent each with “war” and “wirk-
lich”, but the first one would be limited to a bitonal representation L+H* and
thus not map the rise-fall-rise pattern. Postulating an intermediate phrase is
counter-factual in the absence of prosodic phrase boundary signals, and it
would not even solve the problem because the adequate representation of the
pattern requires a fourth pitch point. Only an intonation phrase boundary inside
the unitary utterance would make it theoretically possible to represent the rise-
fall-rise contour within AM/ToBI, but that is completely at odds with the em-
pirical facts, since it conflates two clearly separate patterns.

KIM/PROLAB handle the phrase-internal intonations of “War oder das wirk-
llich?” with the same theoretical stance and with reference to the same repeto-
ire of global pitch categories as the phrase-final ones of “Du.?!”: peak, val-
ley and peak-valley contours, symbolized as <&>, &<&> and <&.>,<&.>,
respectively. The <&> and &< referred to two rising categories with a lower and a higher end point, of which the latter only occurs before a prosodic phrase boundary. The contour types are attached to separately
determined accents, which are default, partially reduced or reinforced, marked by <&2>, &<1>, &<3> before the accented word; unaccented words are
<&0>. Phrasing boundaries are <PGn>, n indicating the strength of separa-
tion. Thus we get:

(4) (a) one-accent case with rise-fall-rise
   <&2> war er das wirklich <&.> <PGn> (<&2> was it really him <&.>
   <PGn>)
(b) two-accent case with rise-fall-rise + rise
   <&2> war er das <&.> <&2> wirklich <&.> <PGn> (<&2> was it him <&.>
   <&2> really <&.> <PGn>)
(c) two-accent case with rise-fall-rise + prosodic boundary + rise
   <&2> war er das <&.> <PGn> <&2> wirklich <&.> <PGn> (<&2> was it him
   <&.> <PGn> <&2> really <&.> <PGn>)
(d) two-accent case with (rise)-fall + rise
&&2 war er das &&. &&2 wirklich &&. &&PGn (<&&2> was it him &&>
&&2 really &&. &&PGn>)

Figure 1: Spectrogram and F0 traces (in Hz) of the utterance “war er das wirklich?”: (a) a single-accent phrase with a rising-falling-rising contour starting on accented “war”, (b) a two-accent phrase with a rising-falling-rising contour on “war er das” and a rising one on “wirklich”, (c) differs from (b) by lengthening of “das” with factor 1.5, introducing a prosodic boundary, (d) a two-accent phrase with a rising-falling contour on “war er das” and a rising one on “wirklich”.

Figure 1 shows the four F0 traces in relation to the spectrogram. F0 was generated in praat from natural (4b), produced by a male speaker (ED) in the Prosody Colloquium: the psola synthesized stimuli are accessible from Kohler (2006, Audio Example 1). Version (4c) differs from (4b) by the duration factor 1.5 being applied to the stretch of speech marked in the F0 tracing (b)(c), corresponding to the word “das”; the rest of the sentence has factor 1.0. This
lengthening produces a clear splitting up into two prosodic phrases, which is not the case in the original stimulus (4b). These examples provide acoustic and auditory illustrations of the theoretical limitations of the tone sequence model in coping with empirical pitch patterns, whereas a pitch contour model can easily handle them. The next chapters will further support this holistic approach with evidence from perception experiments that show the auditory relevance of global features of pitch patterns.

2.2 Synchronization vs alignment

In KIM, contour types are further differentiated by the synchronization of their F0 time courses with the vocal tract timing.

2.2.1 Synchronization of peak contours

A peak contour may be synchronized such that the F0 maximum occurs before or at the accented vowel onset and F0 descends into the accented vowel, across the articulatory landmark of the consonant-vowel transition: early peak. The synchronization may also be inside the accented vowel, with F0 rising across the articulatory landmark: medial peak. The vowel-internal synchronization of the peak contour may be shifted progressively later so that the maximum occurs at the end of the accented vowel or in subsequent unaccented sonorous sounds if there are any: late peak. A very late peak has a low F0 precursor inside the accented vowel. Figure 2 illustrates the three synchronizations in the German sentence “Sie hat ja gelogen.” (“She’s been lying.”); generated in praat from a natural production by a male speaker, KJK), with a single accent, on “gelogen”, around its lexically stressed syllable “lo”. For the early peak the F0 maximum is at the accented vowel onset, for the medial peak it is in the middle of the accented vowel, and for the late peak it occurs at its end. The speech files are in Kohler (2006, Audio Example 2).

These three synchronizations are distinctive categories of the intonational phonology of German and of many other languages, including English. They code different pragmatic meanings, in both German and English, which may be paraphrased as

\[
\begin{align*}
(5) \quad & (a) \text{ early peak } & \text{knowing, coming to the end of an argument, summarizing, resignation} \\
& \quad & \text{finality} \\
& (b) \text{ medial peak } & \text{observing, realising, starting a new argument} \\
& \quad & \text{openness} \\
& (c) \text{ late peak } & \text{observing, realising in contrast to one’s expectation, surprise} \\
& \quad & \text{unexpectedness}
\end{align*}
\]
The perceptual relevance and the pragmatic meaning of the different synchronizations of F0 peak contours with articulation have been analysed in a number of experiments within the KIM framework (Kohler, 1987, 1990, 1991a; Niebuhr, 2003; Niebuhr and Kohler, 2004). The main experiments are the following (further details about experimental procedure in Kohler, 1991a).

Figure 2: The early, medial and late F0 peak contours in the German sentence "Sie hat ja gelogen." synchronized with the onset (a), middle (b) or end (c) of the vowel in the accented syllable "lo", as represented in the speech wave and the spectrogram.

Experiment (A) Starting from a natural production of the German sentence "Sie hat ja gelogen." ("She's been lying."), a series of F0 patterns was generated by shifting the rising-falling F0 contour to the left in 6 equal steps of 30 ms and similarly to the right in 4 equal steps. The LPC parameters of the utter-
ance were then resynthesized with the 11 pitch patterns. The resulting 11 stimuli spanned a continuum from early to late peak synchronizations; Stimulus 5 was the first in the series where the F0 peak maximum was inside the accented vowel. The series was presented in left-to-right peak shift order to 60 listeners, who were asked to indicate whether and at what point(s) in the series they heard changes in the speech melody (serial discrimination test), and what meanings were associated with the changes. Table 1 lists the frequencies of perceptual change responses as they occurred for the stimuli in the series.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st change perceived at</td>
<td>1</td>
<td>4</td>
<td>39</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>further changes perceived at</td>
<td>1</td>
<td>5</td>
<td>11</td>
<td>15</td>
<td>21</td>
<td>22</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>4</td>
<td>40</td>
<td>21</td>
<td>11</td>
<td>15</td>
<td>21</td>
<td>22</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 1: Frequency distribution of ‘perceptual change has occurred’ responses by 60 listeners in the left-to-right sequence of the serial discrimination test across the 11 stimuli with F0 peak shifts in “Sie hat ja gelogen.” (1 = leftmost, 11 = rightmost position).

There is a large and sharp maximum for the first change at stimuli 5/6: 90% of the responses are at this point in the series, when the F0 peak maximum enters the accented vowel and pitch switches from a descending to an ascending pattern across the articulatory landmark of the consonant-vowel transition. The frequencies for further changes in the series are less sharp; they are spread across the remainder of the peak-shift continuum, although there is a minor maximum around stimuli 9/10. These results point to a categorical shift from an early to a medial peak, but to a gradual change from medial to late. As regards the assessment of meaning, there was a wide spread of answers, but they converged on the pragmatic categories knowing, realising, surprised.

**Experiment (B)** With these results, it became possible to construct contexts that verbalise the three pragmatic meaning categories, and to pair early, medial and late peaks with them for judgements as to whether context and stimulus formed a match or not. The contextualizations were the following:

1. (a) **knowing**  “Wer einmal lägt, dem glaubt man nicht, auch wenn er gleich die Wahrheit spricht.” (“Once a liar, always a liar.”)
2. (b) **realising**  “Jetzt versteh ich das erst.” (“Now I understand.”)
3. (c) **surprised**  “Oh”
The 3 x 3 naturally produced pairings of contexts and peak positions were presented in randomized order to 88 listeners for matching judgements in an identification test relating perceived pitch patterns to pragmatic meaning categories (natural stimuli identification test). The results are given in Figure 3.

The early peak fits into the knowing context, where the other two peak synchronizations do not match well, least the late peak, and contrariwise, the early peak is a mismatch in the other two contexts, where the medial and the late peak fit in well. However, the differential bindings of the latter with the realising and surprise contexts, respectively, are not clearly marked. These results, together with the discrimination data of (A), suggest that early and medial peaks are perceptually and pragmatically discretely differentiated, whereas medial and late peaks form a perceptual and a pragmatic continuum.

Figure 3: Percentages of ‘matching’ responses for the combinations of 3 contexts - knowing, realising and surprised - with early, medial and late F0 peaks in the sentence “Sie hat ja gelogen.” in a natural stimuli identification test. 88 subjects.

Experiment (C) The next step was to determine the change-over point from early to medial peak more precisely by a combination of an identification and a pairwise discrimination test, in the framework of categorical speech perception (Liberman et al., 1967).

For the identification test, the context sentence “Jetzt versteh ich das erst.” of (B) was paired with each of the first 8 stimuli of the (A) series of “Sie hat ja gelogen.”, and the contextualized stimuli were repeated 10 times and randomized. To produce the same acoustic quality for context sentence and test stimulus the former was also LPC synthesized (synthesized stimuli identification
19 subjects took part in the identification test and judged the context-stimulus sequence as 'matching'/'not matching'.

For the discrimination test, the 11 stimuli of the (A) series were paired with a step size of 2 in ascending and descending orders. Two test tapes were compiled, one for the ascending and one for the descending order, each containing a randomization of 2 repetitions of all stimulus pairs (randomized paired discrimination test). Two groups judged the pairs as 'same'/'different', 39 subjects in the ascending order, a different group of 34 in the descending order.

Figure 4 gives the results of the identification and discrimination tests. The contextualization of the early to medial F0 peak continuum with the "Jetzt" context shows an abrupt change from 'matching' to 'non-matching' in spite of the gradual change along the physical dimension. Furthermore the paired discrimination test shows a major peak in the discrimination function around stimuli 5/6 (confirming the results of test (A)), beside a strong order effect.

Figure 4: Identification Function (left) in the synthesized stimuli identification test showing percentage 'matching' judgements for 8 stimuli “Sie hat ja gelogen.” with F0 peak shift from left to right in the context “Jetzt verstehe ich das erst.”. 19 subjects, n = 190 for each stimulus. Discrimination Functions (right) in the randomized paired discrimination test, showing percentage of 'different' judgements for utterance pairs of “Sie hat ja gelogen.” with a 2-step distance of F0 peak positions, in the ordering left-right (continuous line) or right-left (broken line). The stimulus numbers refer to the second stimulus in ascending and to the first stimulus in descending order. 39 subjects for left-right ordering, n = 78 at each data point; 34 subjects for right-left, n = 68.
Discrimination is sharpest, in both orderings of stimuli, if the 2-step distance is located around stimuli 5 to 7, but weakens if it is outside this stimulus range. Stimulus 5 is highly discriminated if it comes second or is spanned in the pair, i.e. in 3 – 5, 7 – 5, 4 – 6, 6 – 4, also in first position in the ascending order. The discrimination function for equal stimuli is distinctively different, with false alarms staying well below 50% throughout, but there is an increase of false alarms for Stimulus 5 (Kohler, 1987, 1991a). The results of identification and discrimination together support categorical pitch perception: stimuli 1 – 4 represent one perceptual identification category, stimuli 5 – 8 a different one. They constitute the two phonological categories of *early* and *medial peaks*. Discrimination of stimuli is sharpest between these identification categories, which is what the theory of categorical perception postulates (Repp, 1984).

The reason for this finding can be sought in the switch from a falling to a rising F0 across the articulatory landmark of consonant-to-vowel transition, with its concomitant increase in acoustic energy, at Stimulus 5 in the series. This energy increase heightens the change in pitch movement. Thus, in stimuli 1 – 4, a contrastive high - low pitch movement into the accented vowel is highlighted, whereas in stimuli 5 – 8 it is a contrastive low-high movement.

This discrete perceptual contrast between *early* and *medial peaks* allows the two pitch categories to be associated with the discrete pragmatic categories of *finality vs openness*, which are well supported by the results of Experiment (B). On the other hand, a gradual perceptual change from *medial to late peaks* is mapped onto a graded overlay of *unexpectedness* within the pragmatic category of *openness*. However, the end points of the perceptual and pragmatic scales constitute clearly differentiated categories; they only lack the conformity with the discreteness postulate (cf. 2.2.3/4).

Both *early* and *medial* contour types are peak contours, i.e. are terminal falling, and are perceived as such, and differently from global rising contours. All peak contour synchronizations have this aspect of terminality, but with different weightings of low and high frequencies in the accented vowel. This is further evidence for a holistic treatment of peak contours. At any rate, both early and non-early peaks need to contain a reference to the low ending in their definitions. ToBI, however, relegates the low ending to the phrase accent and boundary tone in the *medial* and *late peaks* (L→H*+L-L% and L*+HL-L%, whereas it is part of the pitch accent in the *early peak* H+L*+L-L%.

2.2.2 External pitch synchronization vs internal pitch timing

The research by Niebuhr (2003) has demonstrated that the perceptual categories of *early* and *medial peaks* are not just determined by the external synchronization of the F0 and articulatory time courses, but also by the internal timing of pitch contours, i.e. the rising and falling speeds in the global peak contours. They may both be slow or fast, or the rise may be slow, the fall fast, or vice
versa. A prototypical early peak has a slow rise and a fast fall; the reverse holds for the prototypical medial peak. The percept of medial peak can thus be created by an early synchronization with a fast rise and a slow fall. This is illustrated by Figure 5 and the speech files in Kohler (2006, Audio Example 3).

Figure 5: The early (a) and medial (d) F0 peak contours in “Sie hat ja gelogen.”, taken from Figure 2, synchronized with the onset and the middle of the vowel in the accented syllable “lo”. The early position is kept in (b) and (c), but in (b) the fall is slowed down, and in (c) the rise is speeded up and the fall slowed down.

The F0 patterns represented in (a) and (d) of Figure 5 are identical with (a) and (b) in Figure 2 and produce the percepts of an early and a medial peak, respectively. If the fall is slowed down, as in (b), the percept is still in the early peak category, but with a softening of its finality. If the slow fall is coupled with a fast rise, as in (c), the percept changes to the medial peak category, and is almost indistinguishable from (d). So there are trading relations in perception between synchronization and contour-internal timing. This is decisive proof
that the listener evaluates intonation contours holistically and that communication between speaker and listener is based on pitch gestalt, which in turn means that intonation must be modelled on tone contours rather than tone-level sequences. In addition, Niebuhr's ongoing PhD research shows that the perceptual globality also includes prosodic aspects that are usually excluded from intonation analysis, viz. acoustic energy and duration of articulatory gestures, both of which enter into a complex network of trading relations with the two F0 timing factors.

2.2.3 The medial-to-late scale

In his ongoing PhD research, Niebuhr has also investigated the relevance of external and internal timing along the scale from medial to late peaks and has found the same interactions between the two timing factors. He has provided evidence that categorical perception does not apply to the medial-to-late scale, although its opposite ends are separate perceptual and pragmatic categories.

Furthermore, following on from Niebuhr's research, the present author is investigating a further subdivision of the continuum from medial to late peaks. First analyses with the semantic differential technique suggest that the scale needs to be subdivided by the introduction of a late medial range (Kohler, 2005). The sentence "Er war mal schlank." ("He used to be slim.") may be realised in three typically different ways, which are illustrated by Figure 6 and the speech files in Kohler (2006, Audio Example 4, produced by the author).

Version (a) is a matter-of-fact statement of an observation/realization, (b) makes this statement contrastively (surprise with acceptance), and (c) adds a personal expressive evaluation to the contrastive statement (surprise with opposing assessment). Version (a) has a medial peak, version (c) is the late peak with a low pitch precursor before the late rise, which is absent in the late medial peak of version (b). Further data can be found in Kohler (2004a).

These contextualizations as well as the results from Kohler (2005) show that there are two late peak categories in the prosodic phonology of German with clear phonetic manifestations and communicative functions. This medial-to-late series forms a continuum inside the accented vowel and differs from the separation of early vs non-early across the articulatory consonant-vowel landmark. Early vs non-early signal finality vs openness, knowing vs realising. As high pitch is successively shifted towards the end of the accented vowel or even beyond, with a concomitant strengthening of low pitch at the beginning, the additional meaning of contrast is superimposed on the meaning of openness. When low pitch is salient enough initially in the accented vowel the speaker's personal expressive assessment comes in as an overlay on the meaning of contrast. Contrast and expressive assessment are graded just like the phonetic manifestations. Nevertheless there are three categories in the series.
medial to late, but there is no categorical perception in the established sense of an abrupt category change with increased discrimination sensitivity.

Figure 6: 3 natural productions of “Er war mal schlank.”: F0, spectrogram. (a) medial peak: statement of observation; (b) late medial peak: contrastive statement; (c) late peak: expressive evaluation.
2.2.4 Synchronization of valley contours

Valleys differ from peaks in their pragmatic meaning by signalling consideration for the addressee. In German, the degree of consideration for the listener, and concomitantly the subordination of the speaker, increases with the end point of the F0 rise. Questions, i.e. requests for information or for polarity decisions, are most likely to be associated with a high rise, particularly if the question is not marked syntactically. The interaction of F0 pattern and syntactic form is, however, quite complex (cf. Kohler, 2004b).

In addition to these differences of shape, the intonational phonology of valley contours also has to recognise differences of synchronization, parallel to the ones for peaks. KIM distinguishes an early valley with the F0 minimum occurring before the accented-vowel onset, and a late valley, where it is located inside the accented vowel. Figure 7 illustrates the two synchronizations in the German sentence “und wie ist dein Name?” (“and what’s your name?”) with a single accent, on the word “Name” around its lexically stressed syllable “Na”. The corresponding speech files are in Kohler (2006, Audio Example 5). The two utterances were synthesized in praat on the basis of a natural early valley production (male speaker KJK), using values from natural utterances of both valleys. The patterns code different sub-categories within the pragmatics of valleys, which may be paraphrased as:

(7) (a) early valley matter-of-fact, casual, little involvement on the part of the speaker
(b) late valley interested, showing the speaker’s feeling and involvement

![Figure 7: Spectrogram and F0 traces (in Hz) of early (a) and late (b) F0 valley contours in the German sentence “Und wie ist dein Name?”](image)

Vertical line = segment boundary [n] – [a:].
The perceptual and pragmatic categorizations of early and late valleys were investigated in the following experiment.

**Experiment (D)** Complementary to the peak shift Experiment (C), the same combination of an identification and a pairwise discrimination test was applied to an early-to-late valley continuum (Niebuhr and Kohler, 2004). In order to obtain comparable results for peak and valley shifts, two stimulus series were generated using the same sentence “Und wie ist dein Name?” The peak and valley patterns were created as mirror images of each other, and the same experimental procedure was applied to both series. A medial peak and a late valley stimulus (accent on “Name”) were produced naturally, each preceded by the context “aha” (“OK”), with a medial peak and a late valley, respectively (male speaker KJK). The communicative particle “aha” with the appropriate intonation provides the setting for ‘+/--matching’ judgements.

F0 manipulation and resynthesis were carried out in praat (cf. Figure 8). The peak maximum/valley minimum F0 was located at the boundary between [n] and [a:] of the accented syllable, and then shifted left and right in 4 and 5 equal steps of 25ms, respectively. The two basic stimuli were then resynthesized with the peak and valley shift continua, creating 10 stimuli each.

![Figure 8: Schematic illustration of the peak and valley contours and their shifts.](image)

The sections from point 2 to 4 were shifted as wholes. Point 3 = 150/84 Hz, point 5 = 70/220 Hz.

For the discrimination tests, each stimulus in the peak or the valley continuum, respectively, was paired with the stimulus two removed in ascending order. In addition, the odd-numbered stimuli were combined to 5 equal pairs. The test procedure was the same as for Experiment (C). For the two identification tests, each stimulus in the series was preceded by the constant context phrase “aha”, either with a medial peak or with a late valley. The medial peak in “aha” expresses pleasant surprise about new information received from the hearer. For a match, these pragmatic features must be followed up in the next phrase “und wie ist dein Name?” This can only be done by another medial peak. The early
peak, which may be paraphrased as "so let's have your name and be finished with it", is not compatible with this context. Analogously, the late valley in "ahag" expresses friendly concern for the hearer, which sounds odd when it is followed by an early valley, expressing casualness. So a late valley in the stimulus is necessary for judging both phrases as 'matching'.

18 subjects took part in the experiment. Stimuli in both test types were repeated 5 times and randomized. The results are given in Figures 9 and 10.

Figure 9: Identification Functions for peak and valley stimuli 1 - 10, 18 subjects, n=90

Figure 10: Discrimination Functions for peak and valley stimuli for unequal pairs 1_3, 2_4 ... 7_9 (connected points) and equal pairs 1_1, 3_3 ... 9_9 (isolated points), 18 subjects, n=90
2.2.5 Phonetic alignment

Synchronization of pitch contours in KIM differs from tonal alignment in the AM framework. The former refers to timing categories of intonational phonology, which are established through experimental prosodic analysis. The latter looks at temporal phonetic manifestations of preestablished timeless phonological categories. A considerable number of investigations into F0 alignment in production have been carried out with this Laboratory Phonology approach, pursuing two goals:

(a) to see whether F0 target points (local F0 maxima and minima) are consistently aligned relative to identifiable landmarks in the segmental string (e.g. syllable and segment boundaries), constituting segmental anchoring of the phonetic manifestations of phonological tones (e.g. Arvaniti et al., 1998; Ladd et al., 1999; Ladd et al., 2000).

(b) to see whether the same pitch accent is aligned differently in different languages or in different dialects of the same language, including the question of truncation vs compression (e.g. Grabe, 1998; Atterer and Ladd, 2004).

The majority of alignment studies deal with the (L+)*H and L*+H tones in prenuclear position, or, followed by L-L%, in nuclear position. In the prenuclear case, pitch accent concatenation may result in a hat or a dapped contour. All these patterns are either medial or late peaks in KIM categorization, i.e. early peaks (H+L*) are not included. None of these analyses establish beforehand and independently of the alignment measurements how many prosodic categories that differ in phonological synchronization are needed for the particular language(s) or dialect(s) under investigation. Nor is it ascertained, by auditory evaluation preceding measurement, whether all the subjects realised the same distinctive tonal pattern(s), and whether, in language comparisons, the patterns are actually the “same” from the point of view of the prosodic systems.

Atterer and Ladd (2004, p. 193) push this methodological problem aside by saying: “...the evidence presented here suggests that the ordinary accentual rise used spontaneously by speakers of English, Dutch, German and Greek on the first accented word of read sentences is in some way the “same thing” cross-linguistically.” “In some way” is completely undefined, and so is the “same thing”. What sort of a “thing” is it epistemologically that the authors are dealing with here? They postulate a rising LH contour, exclude any pattern that does not conform to it in their data base and start measuring. Their results then prompt them to say that it must somehow be the “same thing” that is aligned later in German than in English, and later in Southern than in Northern German. This is an argumentative circle, lacking explanatory power.

The decisive question is how many LH patterns need to be differentiated in the phonology of a language: we need three for German, and likewise for English. The next question is how language differences of earlier or later alignment are integrated into such a system of oppositions of synchronization. Are
the alignments of ALL the distinctive categories shifted in the phonetic realizations of the compared languages? Or does data collection capture different categories for the different languages? Or do speakers of the different languages use the categories with different frequencies and do they therefore produce different prototypes in prototypical reading tasks?

This problem of “sameness” of prosodic category across languages is highlighted by Grabe (1998). Laboratory production data were collected for H*+L in parallel contexts for English and German. It is postulated, but not explicated, that both languages have the same intonational category in their phonological inventories, and that the same category is elicited in the data collection. This category is then filled with phonetic substance through measuring phonetic alignment. The conclusion is that alignment is later in German than in English, and German applies truncation, whereas English uses compression. Grabe’s work is one of the few examples in the Laboratory Phonology framework that have based data collection on contextualization. However, closer inspection reveals that the contexts in the data acquisition are not identical for the two languages. They are as follows:

(8) (a) “Anna and Peter are watching TV. A photograph of this week’s National Lottery winner appears. Anna says: Look Peter! It’s Mr. Sheaf/Shaf/Shift! Our new neighbour!”

(b) “Anna und Peter sehen fern. Ein Lottogewinner wird vorgestellt. Anna sagt: „Na sowa! Das ist doch Herr Schiefer/Schief/Schiff! Unser neuer Nachbar!”

The German context suggests expressively evaluated contrast with “Na sowas!” (‘Well I never!’), reinforced by the contrastive particle “doch”. This connotation is absent from “Look Peter!” in the English context. In such a context, German uses a semantically and pragmatically contrastive late peak. It is thus most likely that Grabe’s German speakers produced a late, her English speakers a medial peak, and it is therefore to be expected that alignment is later in the German than in the English data. Truncation then follows naturally from the greater time constraint in the voiceless context of monosyllables in German than in English. This shows that communicative function is already important at the data collection stage, and that different phonological synchronizations of F0 patterns with articulation need to be distinguished from variable phonetic alignment to avoid misinterpretation.

Grabe’s results have not proved that German aligns H*+L later than English, nor that it is a truncating language, compared with English, which compresses falling pitch patterns. They rather suggest that a different systemic pattern was selected from comparable inventories by the speakers of the two languages through different contextualizations, and that the general phonetic processes of microprosodic interaction operated on the different macro-prosodic base. Taking the elevation of F0 after voiceless obstruents and the
later F0 maximum of a *late peak* into account explains the F0 tracings Grabe found. The almost flat F0 in "Schiff" follows directly from these conditions. There is actually a slight F0 decrease at the end of the accented vowel in the figure provided; it is sufficient to signal to a listener that the pattern is not level but a *late peak*. The perceptual relevance of such a small pitch change for pattern recognition before voiceless obstruents is illustrated by an example from *The Kiel Corpus of Spontaneous Speech* (IPDS, 1995, 1996, 1997) in Figure 11. The corresponding speech files are in Kohler (2006, Audio Example 6).

![Figure 11: Spectrogram and F0 trace (in Hz) of “Besuch in Stockholm bei der ICPhS” with late medial peaks in accented [ølm] of “Stockholm” and accented [øs] of “ICPhS”. The Kiel Corpus of Spontaneous Speech, file g105a000, female speaker UTB. Vertical lines mark the 2 peak contours.](image)

In the German phrase "Besuch in Stockholm bei der ICPhS" ("visit to Stockholm to the (international) C(ongress) (of) Ph(tonic) S(ciences)") from file g105a000, female speaker UTB focuses both the place of the visit and the meeting within it, produces two prosodic phrases, separating them by a pause, and uses *late medial peaks* on the second syllable of accented "Stockholm" and on the last syllable [øs] of accented "ICPhS". But whereas the former rises to 380 Hz and then falls to 215 Hz over quite a long stretch of voicing in [ølm], the latter has to be squeezed into a short vowel before a voiceless fricative and gets its descent curtailed due to lack of time: it falls from a comparable value of 372 Hz by only 15 Hz, i.e., < 1 st, over 30 ms. But this small F0 decrease in a short time interval is sufficient for a listener to perceive the same *late medial peak* as in "Stockholm". This perception is, however, tied to articulated speech
where the listener takes the voiceless segmental context into consideration. In listening to the F0 time course, dissociated from articulation (e.g. hum or pulse contours in praat), the second pattern is different from the first, ending high.

The conclusion to be drawn from this discussion is not that phonetic alignment differences between languages need to be rejected altogether but that a more sophisticated methodology has to be applied in data acquisition and analysis to separate phonological synchronization and phonetic alignment.

2.3 Lab speech vs spontaneous speech

The data presented in this paper within the KIM framework have all been collected and analysed by systematic experimentation in the laboratory. But through the contextualization of stimuli, communicative function has always been part of the investigation, supported by an experimental paradigm which, prior to measurement, takes its starting point from impressionistic observation of auditory and semantic differences between pitch patterns in connected speech, text reading as well as spontaneous interaction. This communicative environment is in turn the basis for setting up hypotheses and experimental designs. In a large number of prosodic studies within Laboratory Phonology (e.g. Atterer and Ladd, 2004), this communicative aspect is lacking, and mechanistically constructed and reproduced sentences, outside a context of situation, are subjected to measurement right away.

Moreover, laboratory data by themselves cannot elucidate the intricate substance-function relationships of natural speech interaction. They cannot capture and explain the use of types of prosodic patterns in types of context of situation. Therefore, the experimental approach requires the complement of corpus analysis of spontaneous speech. The categories developed by the function-oriented experimental paradigm of KIM are linked to speech communication, and so provide an adequate basis for labelling and analysing spontaneous speech data. In the course of this labelling, the intonation model has been expanded and adjusted: the result is the prosodically labelled Kiel Corpus of Spontaneous Speech for German (IPDS, 1995, 1996, 1997). This corpus is the basis for the study of frequency distributions of the various pitch accent and phrase boundary patterns and of their semantics and pragmatics in spontaneous interactions. Search scripts allow the retrieval of all the tokens of the different categories via their symbolizations and their subsequent acoustic analyses, followed by statistical evaluation.

Two questions have been pursued so far: (a) the use of peak and valley patterns in lexical and word order questions and their semantic and pragmatic functions (Peters et al., 2005a; Kohler, 2004b), (b) the use of bundles of prosodic features (pitch patterns, timing, pauses, voice qualities) to give varying strength to the separation of prosodic phrases, including dysfluencies (Peters et al., 2005b; Kohler et al., 2001). In both cases, the descriptive symbol and sig-
nal analysis of corpus data is subsequently used for systematic manipulation of acoustic parameters to investigate the effects on perceptual and semantic change. As regards (a), the patterns found in the corpus are changed to their falling or rising counterparts, and the semantic and pragmatic effects in the corpus contexts are assessed. In the case of (b), the different parameters are systematically varied along physical scales and tested for their separation strength in formal perception experiments.

Thus, the research methodology within the KIM framework aims at integrating corpus analyses and laboratory experiments, with cyclical moves between the two. The function-oriented perspective based on contextualization of utterances combines the two approaches, both in perception and production.

3 Function-Oriented Experimental Phonetics

The foregoing discussion points to the need for a new paradigm in prosodic research beyond the dominant Laboratory Phonology approach. In addition to the modelling of pitch patterns as contours instead of tone-level sequences, the new model will have to rest on three pillars: Function, Time and the Listener.

3.1 Function

Ladd (1996) defines the ‘Linguist’s Theory of Intonational Meaning’ as a view the central idea of which is “that the elements of intonation have meaning. These meanings are very general, but they are part of a system with rich interpretative pragmatics, which give rise to very specific and often quite vivid nuances in specific contexts” (pp. 39f). But “the most basic task of any phonological analysis ... is to identify the categorical or quantal elements in a phonological system and to account for the ways in which the realizations of these elements vary. If intonational and paralinguistic messages are indeed distinct, then one of the sources of variation in the realization of intonational categories is paralinguistic modification, and the basic task of analysing intonational phonology is to tell intonation and paralanguage apart” (pp. 38f).

The components of this theoretical message are:

(9) (a) The linguistically interesting elements are discrete and categorical; this also applies to intonation.
(b) The categories of intonation need to be separated from their paralinguistic variability.
(c) Only these intonational categories are the linguist’s concern.

This is the typical linguistic paradigm, which is also represented in various publications by Gussenhoven (e.g. Gussenhoven, this volume), and in the last resort goes back to de Saussure. It is falsifiable on all counts.
(a) The axiom of categoriality of prosodic elements is falsified in both directions. On the one hand, elements that must be given linguistic status are not categorical. This has been shown in Kohler (2004c) for the differentiation of continuation and question rises in the German sentence “Alle Jungen spielen Fußball.” Although the opposite ends of a scale of low to high rising pitch are clearly differentiated as perceptual and illocutionary categories, there is no quantal jump. On the other hand, the data on the synchronization of early vs medial peaks in 2.2.1 show clear categoriality, although they would be paralinguistic rather than intonational phenomena, under the above definition.

(b) In real-life interactions, linguistic and paralinguistic phenomena are closely interwoven, and no operational criteria have been provided to separate them in natural utterances. Linguists apply a linguistic methodology to deal with this problem. To filter out paralinguistic elements right from the start, they work out language examples, usually of sentence length (and often of dubious semantic and pragmatic validity, as illustrated in Section 1), at their desks and then have them read under laboratory conditions without providing contexts of situation. Under such artificial stylization, extensive paralinguistic expression is unlikely to surface, but on the other hand, the results deduced from such data are not generalizable to communicative interaction in the language under investigation. Such extrapolation is, however, wide-spread in Laboratory Phonology, reifying findings as cognitive categories underlying speech production. There is, of course, no proof from independent experiment that the same prosodic categories underlie all speech production as cognitive entities. This is simply assumed, and since reading tasks of stylized sentences are regarded as being virtually unaffected by paralanguage, the extrapolation is considered justified. But this is an argumentative circle, and the assumption may be refuted by pointing out that reading isolated sentences under laboratory conditions has the paralinguistic features of that reading style, and therefore the disentangling of “intonational and paralinguistic messages” still applies. This comes under what Ladd himself calls ‘paralinguistic stalemates’ (p. 39).

(c) Only a very small part of the meaning conveyed by pitch patterns falls into the linguistic domain, as defined: intonation in connection with accentuation and focus, with prosodic phrasing of utterances, especially in close link with syntactic structure, and, to a very limited extent, with sentence modality. Everything else is paralinguistic according to the definition given: expressive and attitudinal meaning, interactive meaning signalling speaker-hearer relations, meaning that conveys the speakers evaluation of events (finality openness unexpectedness), meaning of speaking style (reading, spontaneous speech in various contexts of situation). The ‘Linguist’s Theory of Intonational Meaning’ is rather an impoverished view of intonation. It leaves out most of what constitutes prosody in speech communication and what is central to the behavioural sciences — to psychologists, sociologists, social anthropologists, as
well as to linguists in conversation analysis (Couper-Kuhlen and Selting, 1996).

The remedy for this dilemma is a new paradigm in prosodic research which, in addition to the linguistic function in speech communication, considers the whole spectrum of communicative functions - expressive, attitudinal, interactive, evaluative and stylistic. This means giving up the ‘Linguist’s Theory of Intonational Meaning’ in favour of looking at prosodic exponents of the whole spectrum of communicative meaning and its interrelations. The prosodic analysis of spontaneous speech is thus no longer put into the Procrustean bed of narrow linguistic categories, but handled by a behaviourally oriented substance-meaning relationship. The prerequisite to this approach is data collection in speech production and perception within contexts of situation created by contextualization.

In this new approach it becomes possible again to search for direct links between phonetic substance and behavioural function. Ohala’s (1983, 1984) frequency code is such an attempt to relate phonetic substance – high vs low F0 – to social behaviour – subordination vs dominance, and to subsequently explain linguistic form with reference to this relationship, e.g. the use of high F0 components (high register, high F0 peaks, rising F0) in questions in the languages of the world. Another instance of the frequency code may be the use of early vs medial peaks in German. As described in 2.2.1, early peak contours focus on low pitch in a high-low transition into the accented vowel, whereas medial peak contours focus on high pitch in a low-high transition. The association of this low vs high pitch with finality vs openness may be seen as another aspect of pitch related to dominance vs subordination, and thus to a very general principle of human behaviour.

3.2 Time

In Section 2.2, it has been shown with reference to experimental perception data that two timing factors determine peak and valley categories – external synchronization with vocal tract timing and internal shaping of the falls and rises in the peak and valley contours. Time is therefore a constituent of these prosodic categories and needs to be introduced in their definitions. This has been accomplished in KIM, and parallels the introduction of time in articulatory phonology by Brownman and Goldstein (1992). It has also been argued that timing for phonological categorization has to be distinguished from phonetic alignment. It has further become clear that phonological timing in peak contours shows different perceptual effects from phonological timing in valleys, and that it is only the difference of synchronization between early and medial peaks that is categorical in the classical sense of abrupt category change together with increased perceptual sensitivity across the category boundaries. In
all other cases, there is category formation between different sections of the F0 scale, but no increase of discrimination.

Kohler (2004c) related this difference to a psychophonetic principle.

(10) (a) Syntagmatic pitch contrasts high-low vs low-high are anchored on the consonant-vowel transition in the early vs medial peak, and are auditorily enhanced (Diehl, 1991) across this articulatory landmark.

(b) This anchoring is absent from the differently timed syntagmatic low-high contrasts for medial vs late (medial) peaks.

(c) The anchoring is also absent from early vs late valleys and from phrase-final low vs high rises, where the decisive pitch differences occur in the accented vowel and later.

In all cases where pitch patterns are not defined as syntagmatic pitch contrasts in relation to syntagmatic articulatory transitions but as pitch characteristics of syllabic nuclei and phrase-final position, the psychophonetic principle does not seem to operate in pitch perception, hence the negative discrimination results in medial to late peak shifts, in valley synchronization and in the height of phrase-final rises, in spite of functional category formation in identification tasks. To this list may be added the discrimination of peak height for emphasis (Ladd and Morton, 1997). Categorical perception in the classical sense is therefore a special case and not essential for pitch categorization.

3.3 The listener

The listener enters the new paradigm in three ways:

(11) (a) as the trained listener in the metalinguistic tasks of labelling and auditory preprocessing of data before instrumental measurement

(b) as listener-orientation of the speaker in speech communication

(c) as the perceiver in speech communication.

All three facets are taken into account in KIM and in data analysis with this model. Special attention is given to the communicative requirement that measured production features are only of value in speech interaction if they can be perceived and further processed by a listener. This means that the collection of speech production data must be complemented by speech perception experiments, which underscore the communicative orientation of the new paradigm.

4 Outlook

Analysis of prosody and intonation is not just a metalinguistic academic pursuit but aims at describing and explaining languages and speech behaviour in
realistic communicative situations. Therefore, the prevalent paradigm of the ‘Linguist’s Theory of Intonational Meaning’ within the Laboratory Phonology framework needs amending by the central concepts outlined in this paper, especially Function, Time and the Listener, with reference to general principles of speech production and perception. The systematic modelling of pitch timing, as synchronization with vocal-tract timing and as contour-internal timing, needs to be applied to production and perception in a large number of diverse languages, with reflections on human speech communication in general. For example, language data on question vs declarative in Bulgarian, Hungarian, Italian, and Russian (d’Imperio, 2000; Gösi and Terken, 1994; Kohler, 2003) can all be explained with reference to these timing categories and the general principles of the frequency code and auditory enhancement.

There has been growing unease with mainstream prosodic theory and practice, especially with ToBI, and there was a fair amount of rumbling at Speech Prosody 2002. This movement will gather momentum in years to come, and the categories discussed in this paper will play an important role for the development of a comprehensive theory of speech communication and for the description of speech behaviour in the languages of the world within a paradigm of function-oriented phonetics.

References


