Beyond Laboratory Phonology
The Phonetics of Speech Communication

Klaus J. Kohler
Institute of Phonetics and Digital Speech Processing
University of Kiel, Kiel, Germany

Abstract

This paper critically examines the practices of laboratory phonology, regarding two issues that have received extensive analysis in a framework of 'phonology-going-into-the-lab': neutralization of voicing and f0 alignment. A plea is made for a new paradigm of 'phonology-coming-out-of-the-lab', built on function, time, and the listener as fundamental pillars of speech communication.

1 Introduction: paradigms in speech research

Every science develops paradigms, sets of theoretical and methodological principles, which are only partly determined by scientific phenomena, far more so by the sociology of science. They are passed on through teaching and an influential, often missionary discipleship, and are finally codified in textbooks and degree curricula. Changing them is the equivalent of a scientific revolution (Kuhn 1970).

Looking at the history of linguistics of the last 125 years we can make out three such scientific revolutions in the study of language:

- the positivist codification of 'historical linguistics' by the Neogrammarians at the end of the 19th c. (Paul 1880)
- the behaviourist codification of 'structural linguistics' in the first half of the 20th c. (Bloomfield 1933)
- the mentalist codification of 'generative linguistics' in the second half of the 20th c. (Chomsky 1957).

They are also reflected in the analysis of the spoken medium. Within the framework of historical linguistics, phonetics developed two strands – descriptive and experimental. Sievers' introduction to descriptive phonetics (Sievers 1876) was the first volume of the series *Bibliothek Indogermanischer Grammatiken* under the editorship of a Neogrammarian team, Delbrück and Leskien among others. It had the subtitle *'for the introduction to the study of the phonetics of Indo-European languages'*, i.e., it was to lay the foundations for the evaluation of sound change in the historical development of languages.

On the other hand, Rousselot (1897-1901) approached the study of the sounds of languages through instrumental and experimental analysis of the physics of speech production, but again with the aim to elucidate sound change, since he also applied his techniques to the modifications of pronunciation in the dialect of a family in order to deduce general sound laws (Rousselot 1892). It was the positivist conviction that the minute study of sound curves and the numbers derived from them provide the objective basis of speech sounds, which cannot be captured by subjective auditory impressions and their description. So instrumental measurement was regarded to unfold the truth. Thus from its inception in the work of Rousselot, and later of Scripture (1902) and Panconcelli-Calzia (1948), experimental, signal-oriented phonetics was conceptualized as a science discipline dissociated from, and in opposition to, the symbol-oriented, descriptive framework represented at the turn of the 19th c. by, among others, Jespersen, Passy, Sievers, Sweet, Viëtor.

Following on from de Saussure (1915), the rise of phonology in the Prague Circle (Trubetzkoy 1939) and in American structuralism (e.g., Bloch 1948) was the systemic answer of the humanities to the focus on atomistic measurement, which had finally dissolved all linguistic categories. Phonology grew into a new discipline within linguistics, establishing a science – humanities dualism in the study of speech: phonetics provides detailed measurements of physical speech parameters, phonology works out the distinctive features of the sound systems of languages. Basically, this dichotomy has stayed with us ever since.

Research groups in psychology and engineering environments, such as Haskins Laboratories, imported linguistic concepts, e.g., the phoneme, and filled them with phonetic substance in production and perception experiments. The adequacy of the linguistic concepts for these investigations outside linguistics was not questioned but was taken as factual. This led to 'categorical speech perception', 'the speech code', and 'the motor theory of speech perception', all anchored on the given linguistic category of the phoneme. The analysis of an array of phonetic parameters, e.g., to differentiate voiced and voiceless plosive phonemes,

gained momentum in production and perception studies within word phonology, i.e., phonologically contrastive words (or word-like logatomes) were investigated in isolation or in metalinguistic frames. It became obvious very quickly that the results were of limited value as regards the explanation of speech communication. In listening experiments, for example, subjects were able to learn to pay attention to fine detail, without thereby establishing communicative speech sound categories.

Within the framework of generative linguistics (Chomsky and Halle 1968), phonology lost its independent status as a level of linguistic analysis and was incorporated into the grammar. This entailed a change of view on the phonologization of morphophonemic alternations, such as the neutralization of voicing of obstruents in morpheme-final position (e.g., German ***Kind*** [t] 'child' vs ***Kinder*** [d] 'children'). Loss of contrast was recognized as a surface-level phenomenon in final position, but, being accompanied by the maintenance of a phonetic difference in the non-final form of the morpheme, the latter was taken to be the *phonological* representation generally.

So, when researchers started taking these new phonological solutions into the laboratory, it was a natural consequence to raise the question as to whether the postulated non-neutralization in phonology was supported by a phonetic contrast, which had not been detected previously due to missing measurements. Attention to phonetic detail of obstruent voicing, in a number of languages that had been reported as having morpho-phonologically conditioned neutralization, resulted in the rejection of the well-established neutralization phenomena, e.g., in German. In the main part of this paper, these analyses will be evaluated, with regard to the methodology of subject selection, word material, and experimental design.

The essence of this research strategy is that the linguistic form of phonological solutions, arrived at outside the laboratory, is theoretically primary, its phonetic substantiation secondary. It may therefore be called the paradigm of 'phonology-going-into-the-lab', which became the methodological core for the Laboratory Phonology conference series and the theoretical pivot of the ensuing conceptualization of laboratory phonology as a natural science (Pierrehumbert et al. 2000). Filling phonological categories with phonetic substance is intended to alleviate the modularization into phonetics and phonology.

The blending of phonology, with its discrete categories, and of phonetics, with its scalar variability, into the laboratory phonology paradigm has led to a new dilemma: phonological categories may now turn out to be gradient. To ascertain their category status inferential statistics comes into play. Statistical significance between sets of measurements is the decisive factor. In perception experiments, attribution to categories is made dependent on categorical perception, which, in its strongest form à la Haskins, requires a sharp change in the identification function coinciding with a sharp increase in the discrimination function.

But this type of phonetic research still does not tell us very much about speech communication. Measurements of signal parameters are related to phonological categories which are established by techniques that result from phonological theories. But neither phonological categories nor phonetic measurements and their statistical evaluation need necessarily represent language structures that are relevant for communicating individuals, and they may even represent incongruous metalinguistic domains. So extrapolation of such experimental results to categories for real speakers and listeners is problematic, although standard practice in laboratory phonology. In summary, we can say that this type of classic laboratory phonology, in spite of its sophisticated theorising and analysis techniques, and although it has contributed a great deal towards a bridging of the phonetics – phonology schism, has not overcome the philosophy of science approach of 20th c. speech analysis. A new paradigm is needed which has the elucidation of speech communication as its goal and therefore subordinates all phonological categories and phonetic measurements to the relevance for communicative functions. In this approach, phonological form is not primary

but results from experimentally relating phonetic substance to independently motivated communicative functions, including more narrowly linguistic functions. This is the paradigm of 'phonology-coming-out-of-the-lab' or of experimental phonology (Ohala and Jaeger 1986, Kohler 2005a).

To illustrate some moot points in the basic tenet of laboratory phonology, two issues will be discussed, one from segmentals, the other from prosody, which have received extensive analysis in the laboratory: (1) the neutralization of voicing in word-final position, with special reference to German, (2) f0 alignment in intonational phonology. I am well aware that the proponents of laboratory phonology maintain that "laboratory phonology is not a framework...it is a coalition amongst groups of people, with some working in one or another of the various current frameworks, and others working in no phonological framework at all." (Pierrehumbert et al. 2000: 279). But the "hallmark of a successful community is maintenance of a common vocabulary – which can be used by opposing parties in an argument – even at the expense of a gradual drift in both the meanings of technical terms and the empirical domain under discussion." (ibid.: 276). This stance clearly refers to what I have called the sociology of science. Although there are researchers working in this community of laboratory phonology who have moved towards the paradigm of experimental phonology referred to above, there is no denying the fact that 'phonology-going-into-the-lab' has remained the prevalent core of research.

The discussion of the two areas of laboratory phonology investigations will be followed by some thoughts on categories and methods that will have to be central in a new paradigm for the phonetics of speech communication.

2 The issue of neutralization of voicing in final obstruents

It has been an established statement of fact in German philology for at least a century that there is neutralization of word-final lenis with fortis obstruents ('Auslautverhärtung') in German. Trubetzkoy (1962: 213) refers to Standard German (besides 'Russian, Polish, Czech etc.') in his discussion of the cancellation of a phonological opposition at the boundaries of words and morphemes, more particularly as regards final devoicing. Furthermore, it is well-known from foreign language teaching that native German speakers have great problems in acquiring a word-final voiced – voiceless opposition in languages, such as English or French, where the distinction is made.

This firm stance on complete neutralization of the word-final voicing contrast in German was called into question in the wake of generative phonology, when morphologically related word forms were given the same underlying phonological representation with the same distinctive phonetic features, and when Dinnsen (1983) maintained that careful phonetic studies would reveal that all declared neutralization rules in the phonologies of languages were in effect not neutralizing at all. An early example of such a phonological reassessment of neutralization is the statement by Malmberg (1943) that lenis voiced and fortis voiceless plosives in French stay different in the force features even when they lose their voice distinction in regressive assimilation. Partial differences in properties such as duration and intraoral pressure were subsequently analyzed in read laboratory speech by Thorsen (1966) and Fischer-Jørgensen (1968a, b).

The first to take the German neutralization issue into the laboratory were Mitleb (1981), Port et al. (1981), and O'Dell and Port (1983), then followed by more extensive accounts and discussions in Port and O'Dell (1985). In these studies, consistent differences were found, especially in the duration of the vowel preceding the final obstruent, in citation form pronunciations of word lists, containing pairs such as *Rad* 'wheel vs *Rat* 'advice', *Bad* 'bath' vs *bat* 'requested', *Alb* 'elf' vs *Alp* 'mountain pasture', *weg* 'away' vs *Weck* 'breakfast roll'

(southern dialect word), *schrag* (imperative form of obsolete *schragen*) 'join wooden poles cross-wise or slanting' vs *schrak* (past tense of intransitive *schrecken*, which only occurs in *erschrecken*) 'had a fright'. The differences were small (15ms of vowel and burst durations, 5ms of closure voicing on average), but in the same direction as in the non-neutralizing word-medial context, i.e., longer vowel, shorter burst, and more closure voicing in the underlying voiced plosive, some of them reaching statistical significance. The production data were supplemented by listener judgments, which identified the produced member of a pair poorly, but significantly better than chance (59% overall). The authors concluded from these results, buttressed by inferential statistics, that German had incomplete neutralization of the underlying word-final voicing opposition in obstruents, and that these phonetic differences would have to be taken into account in the phonology of German by the introduction of implementation rules of the generally maintained phonological contrast.

Fourakis and Iverson (1984) argued convincingly, on the basis of data from orthographic vs free elicitation of relevant words, that the phonetic differences were caused by hypercorrection of isolated words in a reading task, and that there were no grounds for postulating incomplete neutralization in the phonology of German. They also pointed out that *weg* (although etymologically related to the noun paradigm *Weg* : *Wege*) was not part of a morphological alternation at all, now being an adverb with a short vowel. The orthography argument was further supported by Port and Crawford (1989) in an investigation of the influence of speech style. They found that discriminant analysis to classify productions by underlying final voicing was least successful (appr. 55%) when target words were embedded in sentences that did not draw attention to minimal pairs, but most succesful (78%) when speakers dictated the words. The conclusion from these results was, however, that many small acoustic differences are involved in maintaining the underlying distinction, which surfaces in different degrees depending on speech style.

There is another crucial point that has to be considered in the evaluation of these data and in the inferences for phonological theory and for the phonological explanation of voicing oppositions in German. The authors found it necessary to explain the meanings of some of their words to their subjects before the actual tests. This was certainly necessary for *Alb*, *Alp*, *schrag*, *schrak*, *Weck.* In view of the complete lack of context for words, which in addition are not all part of the subjects' common vocabulary or are unusual morphological forms, the test items assumed the status of logatomes. It also needs to be taken into account that the data were collected in the US from German speakers who had spent considerable lengths of time in an English-speaking environment and had no doubt been made aware of the need to differentiate final orthographic <b>, <d>, <g> from <p>, < t>, <k>. In view of these experimental design constraints, there is no avoiding the conclusion that the data resulting from this 'phonology-into-the-lab' paradigm have neither contributed to the elucidation of the phonetics of speech communication in German, nor have they been able to substantiate the claim of incomplete neutralization; they have thus not advanced our understanding of the phonetics – phonology interface.

The amount of time and effort that went into the attempt to substantiate preestablished phonological categories would have been better spent on assessing the variability of German plosives across different communicative functions, paying attention to situational, syntactic, lexical, and phonetic contexts. In pursuit of this line of research, insights can be gained especially from the following topics of investigation:

- glottalization (Kohler 1994, 1999a, b, 2001a)
- gestural reduction (Kohler 2001b)
- lenition, place and manner assimilations (Kohler 2001c).

In these studies, the question is not as to how phonetic data can be fitted into currently proposed linguistic solutions of linear segmental word phonology. Existing phonemic representations of words are not taken as ontologies of cognitive representations, but simply as heuristic devices for accessing labeled acoustic data bases to investigate the structured variability of the phonetic dynamics of connected speech in various communicative settings. The aim is to derive general principles of the organization of production units beyond the phoneme under the whole array of contextual constraints from situation to phonetics. The timing and economy of more global gestural units, such as opening – closing and closing – opening movements, become the focus of attention. This allows us to analyze time courses of articulatory and acoustic phonetic parameters in their relation to the production and perception of words and phrasal structures. This also makes comparison across languages, dialects and diachronic stages possible, and opens new perspectives on historical sound change, more akin to the commonly pooh-poohed atomistic approach of the Neogrammarians than to the systemic conceptualization of historical phonology.

One such phonetic property of production and perception beyond the phoneme is the phenomenon of glottalization The manifestation of plosives as glottal stops or irregular glottal pulsing in nasal environments is very common in German, e.g., *könnten* 'could' vs *können* 'can'. It is no longer seen as a simple segmental replacement but as the most elementary glottal air stream regulation, allowing continued velic lowering and at the same time conveying to a listener the percept of a break, which can be linked to a stop. Glottalization instead of plosive articulation thus constitutes a more economical gesture to achieve a constant purpose for a listener. Moreover, the positioning of irregular glottal vibration within the nasal stretch can be highly variable in production, and for a listener its occurrence somewhere in this consonantal configuration, or even towards the end of the preceding vowel, is sufficient to decode a stop.

Similarly, the assimilation of German unreleased coronal plosives to labial or dorsal ones across word boundaries cuts out a gesture without hampering intelligibility because the auditory effect of a coronal closing gesture intertwined with a following labial/dorsal one is only minimally different from that of a labial/dorsal closing + long hold + opening. So again gestural economy achieves a constant purpose for the listener.

Finally, the realization of voiced approximants instead of intervocalic lenis plosives and of lenis plosives instead of intervocalic fortis ones in German are economical reductions of closing – opening gestures, maintaining the difference in dynamic energy between the two sets of lenis and fortis gestures at lower values, which is sufficient differentiation for the listener.

In all these cases, phonetic data analysis of connected speech in the laboratory contributes to the linguistic modelling of speech production and perception, no longer filling phonological entities with phonetic measurement, but creating phonological structures from phonetic measurement.

Looking at plosive categories from their phonetic manifestations rather than from their positions in phonemic systems also provides a common tertium comparationis across languages. So we can assert that both German and English show the same phenomenon of glottalization, related to plosives in a nasal environment, e.g., in German *könnten* and English *mountain*, at least in American English (Pierrehumbert and Frisch 1997) and non-RP varieties of British English.

Likewise, a force feature differentiating the two plosive classes /b/, /d/, /g/ vs /p/, /t/, /k/ in German, even when lenition reduces them to a distinction between approximants and lenis plosives, can be extrapolated to other languages, e.g., English, as a dynamic phonetic property. But at first sight such a cross-linguistic generalization seems to be falsified by data from American English, where /t/ may be deleted in such words as *winter*, coalescing with

winner, whilst /d/ stays as a plosive in *cinders* etc. This speaks against referring /t/ and /d/ to a difference in force of articulation, because one would expect the weaker member of the opposition to reduce in a common context.

However, if instead of focussing on an individual consonantal segment the whole closing – opening gesture is taken into account, all the superficially disparate observations fall into place. Intersyllabic /t/ vs /d/ is characterized in English generally by /t/ inducing a fast closing movement as against a slow closing movement for /d/. This is an extension of what is found word-finally in, e.g., *bite* vs *bide*. It takes care of the difference in vowel duration of *writer* vs *rider*. If the dyad vowel + stop closure is kept at the same magnitude in both cases the occlusion phases will show complementary duration differences, and this may be supplemented by VOT differences, again stronger for /t/ than for /d/. On the other hand, the faster closing movement for /t/ intervocalically may change the mechanics of the movement and result in a flap, as against a stop for /d/. This relatively subtle difference can then easily be leveled out, and the difference in vowel duration remains the only distinguishing feature. Even this may be overridden, with the result of complete coalescence.

In the vowel + nasal + /t/ or /d/ context we can expect the same basic dynamics. Here the nasal is also shorter before /t/ than before /d/, i.e., the faster tongue movement is coupled with a faster velic raising to reach the dynamic goal, the establishment of a complete supra-glottal enclosure, more quickly. But the tongue tip mechanics are already initiated for the nasal, which may become flapped for intervocalic /nt/. In a further step, *winter* could coalesce with *winner*, but /nd/, having the slower closing movement, would stay different.

So, what looks like a reversal of the hypothesized strength relationship in the contextual reduction of /t/ vs /d/, would be relatable to the same prototypical category manifestations under changes of the whole syllable dynamics. These English data would then all fit into the lenition patterns one finds in German and in many other languages. With this line of argument, a force feature difference in stop productions /p/, /t/, /k/ vs /b/, /d/, /g/ gains further credibility, rather than being undermined. This does of course not mean that the force difference is always maintained in a synchronic or diachronic perspective; there is always the possibility of coalescence across dialects, speaking styles and historical stages of a language, as shown by German 'Auslautverhärtung'.

3 The issue of f0 alignment with articulation

The framework of autosegmental metrical (AM) phonology (cf. Ladd 1996) and its labeling tool ToBI (Beckman et al. 2005) postulate the reduction of phonetic f0 patterns to phonological points in the form of pitch accents, phrase accents and boundary tones, and further to binary oppositions L vs H in each case. For pitch accents, this basic theoretical framework allows four bitonal units L*+H, L+H*, H*+L, H+L*, besides the two single tones L* and H*. No external grounds were ever provided to justify the combinatorial restrictions.

If empirical data related to pitch accents in a language under investigation cannot be represented by any of these categories but require further differentiation there are two ways of dealing with the problem theoretically: (a) The type of tonal categorization in AM phonology is not adequate and the theory needs fundamental revision, or (b) the theory is kept in its existing detail and new ad hoc categories are added that can cope with the new empirical data. The second solution was adopted by Grice (1995) for English, by allowing three tones in pitch accents in order to account for distinctive pitch phenomena in connection with downstep and with the analysis of 'high plus mid accented syllable' as (H+L)+H* vs 'high plus low accented syllable' as H+L* (pp. 219ff; cf. also Ladd 1996, pp. 109ff). There is an argumentative circularity here: the empirical data force her to expand the categorical repertoire, and then phonology explains the empirical data.

Pitch accents are associated with certain stressed syllables, and in the case of bitonal categories their components either have primary (strong) status or they are leading (up to) or trailing (from) the primary component. This led to the star concept. These primary associations of bitonal components (conceptualized outside the time scale) were subsequently subjected to instrumental measurement, within the laboratory phonology framework, to analyze their f0 maximum/minimum alignment with the stressed syllable and thus to put the phonetic manifestation of the phonological categories on the time scale post hoc. It turned out that H* and L* did not always show the relation in the physics of speech as one would expect from tone association to syllables. So the concept of phonetic alignment was born. The question then became 'how is *phonological association* mapped onto *phonetic alignment*?'

As a way out of this dilemma, Atterer and Ladd (2004) propose to negate starred accents altogether in collecting phonetic measurements of LH and HL trajectories across different languages. They assume identical cross-linguistic pitch categories without justification, and maintain that "serious investigation of differences between languages and dialects must be based on instrumental data. Notational distinctions such as L*+H vs. L+H*... are incapable of representing the range of fine phonetic differences that can be discovered instrumentally." (p. 194). This is the reinstatement of the philosophy of science of experimental phonetics at the turn of the 20th c. It is a step back to the early days of instrumental speech analysis and to Scripture's dictum that it would be best for phoneticians to be born deaf because they would never follow subjective impressions but would only be guided by the speech curves and the numbers derived from them, which contain the truth (Scripture 1936: 135). This is obviously no solution, and should not be considered further.

Alternatively, the incommensurability of theory and data is again resolved by the addition of a new concept, viz. secondary association, set apart from primary tone association by secondary alignment of edge tones to stressed syllables, to moras and to word edges (Pierrehumbert and Beckman 1988). This new phonological concept allows to 'explain' phonetic alignment that deviates from what would be expected from primary association. It becomes a deus ex machina, which always fits, and therefore has no explanatory power.

In applying the concept of secondary association to Catalan, Neapolitan and Pisa Italian, Prieto, d'Imperio, and Gili Fivela (forthcoming) expand it so that not only phrase edges can seek edges of metrical units as anchor points, but also tones in pitch accents. For Catalan, they describe a three-way phonological contrast in rising prenuclear accents and a binary contrast in nuclear position. In commands or in narrow focus statements, the peak of a prenuclear rise is aligned with the right edge of the accented syllable, in a broad focus statement, a prenuclear rise has a delayed peak, and in a polar question or a request, the prenuclear rise only starts on the posttonic syllable. As in prenuclear position, the nuclear accent rise in commands has its peak aligned with the end of the accented syllable, whereas in requests it only starts at that point. The threefold alignment contrast in the penuclear accent rise cannot be captured by the dichotomy L+H* vs L*+H, which the AM framework provides for such rises. This clash between phonological association of pitch accents to syllables and phonetic alignment in and across syllables is again resolved by secondary association. There are two pitch accents, L+H* and L*+H, in primary association with the accented syllable. Their primary associations are sufficient to distinguish the rise with delayed peak and the posttonic rise, respectively. The peak alignment with the edge of the syllable is captured by the addition of a secondary association with the metrical anchor of the accented-syllable edge.

In Neapolitan narrow focus vs question, there are two contrastively aligned LH rising accents (L+H* and L*+H), which are nevertheless realised within the stressed syllable boundaries. This means that there is a clash between contrastive earlier vs later alignment and primary association with the same syllable. The introduction of secondary association is to resolve this clash. It associates the starred tone, H* or L*, secondarily to the first mora of the

accented syllable, and thus reestablishes a phonological difference of association, as represented by the different star categorizations, in spite of their association with the same syllable.

The postulate of secondary association to account for the manifestation of broad and narrow focus in Pisa Italian is more complex. Again there are two contrastive LH patterns associated with the stressed syllable (and followed by a low intermediate phrase boundary tone). Both have a phonetic leading tone [L+], narrow focus also has a trailing low tone. Both pitch accents have H* associated with the stressed syllable, but this H* is differently aligned within the syllable. So both have a primary association of H* with the stressed syllable, but differ in secondary association to the right edge of the syllable for broad focus, but with the right edge of the first mora in narrow focus.

There is no theoretical necessity to deal with the data from these Romance languages in this way; the terms mora and secondary association are not defined within the prosodic theory, but are ad hoc constructs to deal with the empirical data in a clash with the initial phonological representation. For one thing, it can be argued that the introduction of the mora or the syllable edge or the word edge as an anchor for the pitch pattern would make both primary and secondary association superfluous, if it were not for the initial postulate that pitch accents are associated with the stressed syllable. Secondly, although the authors base their analyses on functional contrasts, they nevertheless take preestablished formal phonological categorizations within the AM framework into the laboratory and fill them with measurements, i.e., they adhere to the principal tenet of laboratory phonology. However, this methodological approach forces them to introduce new abstract categories to save the initial postulate and at the same time reconcile it with empirical data.

The much simpler theoretical alternative is to introduce time into the phonological pitch categories in their initial definitions; earlier and later *synchronizations* of pitch contours with syllable articulation can then be differentiated as contrastive categories in the prosodic phonology of Catalan and the two Italian varieties, as well as other languages. The inclusion of time in intonational phonology is comparable to its introduction in articulatory phonology (Browman and Goldstein 1992). But to do this we should aim at the first solution mentioned above: AM theory needs to be changed.

This was first proposed for the intonational phonology of German in Kohler (1987), long before the application of AM to German and before the codification of the labeling tool GToBI (Grice and Baumann 2002). 'Time' became established as a central prosodic category in the development of *KIM* - the *Kiel Intonation Model* (Kohler 1990, 1991a, b, 1995, 1997, 2003, 2004a, c, 2005b, 2006; Niebuhr 2003; Niebuhr and Kohler 2004). KIM conceptualizes intonation as consisting of sequences of peak or valley or combined peak-valley patterns, each linked to one accented word. They are global pitch units (rising-falling, falling-rising, rising-falling-rising), not series of target pitch points. The distinctive units are established by function-oriented experimental phonetics, i.e., the phonological categories are not worked out at the linguist's desk and then carried into the laboratory to be filled with substance, but experimental procedures in functional (contextual) settings lead to the categories.

The peak and valley units have different contrastive phonological 'synchronizations' with articulation: 'early' – 'medial' – 'late peaks, and 'early' – 'late valleys'. These categorizations resulted from perceptual experiments with peak and valley shift paradigms testing discrimination, as well as identification through contextualization (Kohler 1987, 1990, 1991a, b; Niebuhr 2003; Niebuhr and Kohler 2004). They have the time dimension in them, it need not be added post hoc to timeless abstract tonal entities associated with syllables. So the problem of mapping phonological association onto phonetic alignment does not arise.

These peak and valley categories are related to linguistic and paralinguistic functions. In German (Kohler 2004b, 2005b, 2006, Niebuhr and Kohler 2004), peaks code

- 'finality' when 'early': *knowing, summarizing, coming to the end of an argument*,
- 'openness' when 'medial': *observing, realising, starting a new argument*,
- 'unexpectedness' when 'late': *observing, realising in contrast to one's expectation, surprise, disbelief.*

Valleys differ from peaks by the pragmatic meaning of addressee orientation, expressing

- 'casualness' when 'early',
- 'friendly concern' when 'late'.

In addition to these external synchronizations, internal timing of pitch patterns, e.g., fast rise – slow fall or slow rise – fast fall in peak contours, contributes to their distinctive functions (Niebuhr 2003). The prototypical 'early peak' may be defined as slow rising before the accented syllable, followed by fast falling into the accented vowel, as against the prototypical 'medial peak' with a fast rise into the accented syllable and a slow fall in the accented vowel. This produces an opposition of a high – low vs a low – high pitch contrast in the f0 trajectory into the accented vowel, perceptually accentuating low vs high pitch in the vowel. In the case of the 'late peak', the syntagmatic fast rise – slow fall contrast is shifted towards the end of the accented syllable.

The same threefold peak synchronizations have been found in Southern British English (Kleber 2005). The Neapolitan question – narrow focus dichotomy also fits into this frame: a late peak with a slow fall intensifies high pitch late in relation to the accented syllable of a question. Similarly, the Pisa data for broad and narrow (contrastive) focus exhibit a basic contrast in phonological synchronization, combined with internal timing differences to accentuate high pitch in broad focus (cf. Kohler 2003). The same considerations apply to Catalan question vs statement, request vs command and broad vs narrow focus.

Within the differently synchronized categories, there is phonetic variability. This is what most alignment studies analysed (Atterer and Ladd 2004, Arvaniti et al. 1998, Ladd et al. 1999, 2000). It is, however, essential to differentiate phonological external and internal timing from phonetic alignment of pitch. In language comparisons, functional timing has to be established first, before phonetic alignment differences can be evaluated. Putting languages on an alignment scale from early to late, such as English, Dutch, North German, South German, Greek (Atterer and Ladd 2004), presupposes reference to the same functional synchronization across the intonational systems. Then it becomes insightful to say that, e.g., medial peaks are aligned later in X vs Y.

4 The new paradigm of experimental phonology

A new paradigm is needed for the analysis of segmental and prosodic aspects of languages.
Its goal is the elucidation of speech communication, based on 'function', 'time', and 'the listener'. Communicative functions, beyond linguistic form, must be at the center of phonetic investigations into speech and language. The analysis is to be based on situational, pragmatic, semantic, syntactic, and phonetic contextualization. Time is central to the paradigm, not just with reference to duration measurements but as the structuring of the time courses of phonation and articulation, and their synchronizations; time enters the definitions of segmental and prosodic units. At the prosodic level, different synchronizations of f0 contours for different communicative categories are differentiated from phonetic alignment viariation within these categories. The listener also assumes a central role, even in production studies, because communication presupposes that speech is produced for a listener and must therefore be decodable by a listener.

We need to give systematic analyses of large corpora of speech interaction in individual languages a more prominent place in speech research. In a first step these data are annotated, in computer-readable form, with language-specific inventories of segmental and prosodic labels, based on provisional phonemic and prosodic representations, not as linguistic truths about languages, but as heuristic devices to access large corpus data systematically. Segmental transcription starts from canonical word forms in phonemic representation, and relates signal portions to them, with symbolic marking of assimilations, elisions, and insertions. *The Kiel Corpus of Spontaneous Speech* (Institute of Phonetics and Digital Speech Processing (IPDS) 1995, 1996, 1997) is such a database.

Context-sensitive search operations then excerpt, via the symbolic labels, all instances of specific sound classes (and their labeled modifications) in specific segmental and prosodic contexts, or of specific pitch patterns, together with the speech signal portions the labels have been allocated to. Measurements are performed on the excerpted sound classes and pitch patterns, and descriptive and inferential statistics are applied to the symbol and signal data (e.g., Kohler 2001b, 2004b).

Based on the results of the corpus analysis, speech data are collected according to systematic experimental designs, in production as well as in perception, with careful consideration of method, as regards selection of subjects and speech material, also replacing the isolated word and sentence metalinguistic levels by contextual and communicative frames. With regard to the application of statistical procedures to the data, inferential significance needs more sophisticated interpretation, and in particular, it is necessary to distinguish between categorical perception and category interpretation, as communicative categories may be available to speaker and listener without being supported by categorical perception (Niebuhr and Kohler 2004). Thus a methodology will be developed to refine experimental techniques by integrating experimental with corpus data analysis (e.g., Kohler 2004b, 2005b).

Context-sensitive search operations, measurements on corpus data and systematic experimental analysis will then lead to a revision of the initial heuristic categories. The result is the paradigm of 'phonology-coming-out-of-the-lab' or experimental phonology. In parallel, this paradigm needs to be strengthened by function-oriented experimental phonetics, dealing with such concepts as the 'frequency code' (Ohala 1983, 1984), 'minimization of effort', 'self-organization in sound systems', 'hypo-hyper variation' (Lindblom 1990), 'auditory enhancement' (Diehl 1991). There are clear signs of movement among the laboratory phonology community in the direction of such a refocussed experimental approach, with research beginning to center more on communicative function and large databases of connected natural speech, especially in the analysis of intonation and rhythm. These individual activities deserve to enjoy being more widely recognised and practised, and developed into a coherent goal: the 'phonetics of speech communication'.

References

Arvaniti, A., Ladd, D. R., and Mennen, I. (1998). 'Stability of tonal alignment: the case of Greek prenuclear accents', *Journal of Phonetics* 26: 3-25.

Atterer, M., and Ladd, D. R. (2004). 'On the phonetics and phonology of "segmental anchoring" of F0: evidence from German', *Journal of Phonetics* 32: 177-197.

Beckman, M. E., Hirschberg, J., and Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework, in Sun-Ah Jun (ed.) *Prosodic Typology. The Phonology of Intonation and Phrasing.* Oxford: Oxford University Press.

Bloch, B. (1948). 'A set of postulates for phonemic analysis', *Language* 24: 3-46.

Bloomfield, L. (1933). *Language*. New York: Henry Holt and Company.

Browman, C. P., and Goldstein, L. (1992). 'Articulatory phonology: an overview', *Phonetica* 49: 155-180.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N., and Halle, M. (1968). *The Sound Pattern of English.* New York: Harper and Row.

de Saussure, F. (1915). *Cours de linguistique générale.* Paris: Payot.

Diehl, R. (1991). 'The role of phonetics within the study of language', *Phonetica* 48: 120-134.

Dinnsen, D. A. (1983). *On the Characterization of Phonological Neutralization*. Bloomington: Indiana University Linguistics Club.

Fischer-Jørgensen, Eli (1968a). 'Voicing, tenseness and aspiration in stop consonants, with special reference to French and Danish', *Annual Report Institute of Phonetics University of Copenhagen (ARIPUC)* 3: 63-114.

_____ (1968b). 'Les occlusives françaises et danoises d'un sujet bilingue', *Word* 24: 112-153.

Fourakis, M., and Iverson, G. K. (1984). 'On the 'incomplete neutralization' of German final obstruents', *Phonetica* 41: 140-149.

Grice, M. (1995). 'Leading tones and downstep in English', *Phonology* 12: 183-233.

Grice, M., and Baumann, S. (2002). 'Deutsche Intonation und GToBI', *Linguistische Berichte* 191: 267-298.

Institute of Phonetics and Digital Speech Processing (IPDS) (1995). *The Kiel Corpus of Spontaneous Speech*, vol. 1, CD-ROM#2. Kiel: IPDS.

_____ (1996). *The Kiel Corpus of Spontaneous Speech,* vol. 2, CD-ROM#3. Kiel: IPDS.

_____ (1997). *The Kiel Corpus of Spontaneous Speech,* vol. 3, CD-ROM#4. Kiel: IPDS.

Kleber, F. (2005). *Experimentalphonetische Untersuchungen zu Form und Funktion fallender Intonationskonturen im Englischen.* Unpublished MA dissertation. University of Kiel.

Kohler, K. J. (1987). 'Categorical pitch perception', *Proceedings of the 11th International Congress of Phonetic Sciences*, vol. 5, Tallinn, 331-333.

_____(1990). 'Macro and micro F0 in the synthesis of intonation', in J. Kingston and M. E. Beckman (eds*), Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech.* Cambridge: Cambridge University Press, 115-138.

_____(1991a). 'Terminal intonation patterns in single-accent utterances of German: phonetics, phonology and semantics', *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)* 25: 115-185.

_____(1991b). 'A model of German intonation', *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)* 25: 295-360.

_____(1994). 'Glottal stops and glottalization in German. Data and theory of connected speech processes', *Phonetica* 51: 38-51.

_____(1995). 'The Kiel Intonation Model (KIM), its Implementation inTTS Synthesis and its Application to the Study of Spontaneous Speech',
URL: http://www.ipds.uni-kiel.de/kjk/forschung/kim.en.html

_____(1997). 'Modelling prosody in spontaneous speech', in Y. Sagisaka, N. Campbell, and N. Higuchi (eds), *Computing Prosody. Computational Models for Processing Spontaneous Speech*. New York: Springer, 187-210.

_____(1999a). 'Articulatory prosodies in German reduced speech'. *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, 89-92.

_____(1999b). 'Plosive-related glottalization phenomena in read and spontaneous speech. A stød in German?' Poster presented at ICPhS 99 Satellite Meeting on Non-modal Vocal-Fold Vibration and Voice Quality, San Francisco, 31 July, 1999.
URL: http://www.ipds.uni-kiel.de/kjk/pub_exx/kongrbtr/plosglot.html

_____(2001a). 'Plosive-related glottalization phenomena in read and spontaneous speech. A stød in German?', in N. Grønnum and J. Rischel (eds), *To Honour Eli Fischer-Jørgensen. Travaux du Cercle Linguistique de Copenhague*, vol. 31. Copenhagen: Reitzel, 174-211.

_____(2001b). 'The investigation of connected speech processes: theory, method, hypotheses and empirical data', *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)* 35: 1-32.

_____(2001c). 'Articulatory dynamics of vowels and consonants in speech communication', *Journal of the International Phonetic Association* 31: 1-16.
Cf. also URL: http://www.ipds.uni-kiel.de/kjk/publikationen/audiobsp.en.html

_____(2003). 'Neglected categories in the modelling of prosody: pitch timing and non-pitch accents', *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 2925-2928.

_____(2004a). 'Wieviele späte Gipfel braucht die Modellierung der Prosodie im Deutschen? – Methodische und theoretische Reflexionen zur prosodischen Phonologie',
URL: http://www.ipds.uni-kiel.de/kjk/forschung/lautmuster.en.html , section 4(b).

_____(2004b). 'Pragmatic and attitudinal meanings of pitch patterns in German syntactically marked questions', in G. Fant, H. Fujisaki, J. Cao, and Y. Xu (eds*), From Traditional Phonology to Modern Speech Processing. Festschrift for Professor Wu Zongji's 95th Birthday*. Beijing: Foreign Language Teaching and Research Press, 205-214.
Cf. also URL: http://www.ipds.uni-kiel.de/kjk/publikationen/audiobsp.en.html

_____(2004c). 'Categorical speech perception revisited', *Proceedings of the Conference "From Sound to Sense: 50+ years of discoveries in speech communication"*. MIT Cambridge, 11 - 13 June 2004.
Cf. also URL: http://www.ipds.uni-kiel.de/kjk/forschung/lautmuster.en.html , section 4(b).

_____(ed.) (2005a). *Progress in Experimental Phonology. From Communicative Function to Phonetic Substance and Vice Versa*. *Phonetica* 62.

_____(2005b). 'Timing and communicative functions of pitch contours', *Phonetica* 62: 88-105

_____(2006). 'Paradigms in experimental prosodic analysis. From measurement to function', in S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, and J. Schließer (eds), *Methods in Empirical Prosody Research.* (= Language, Context, and Cognition, 3) Berlin/New York: de Gruyter, 123-152.
Cf. also URL: http://www.ipds.uni-kiel.de/kjk/publikationen/audiobsp.en.html

Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press.

Ladd, D. R. (1996). *Intonational Phonology.* Cambridge: Cambridge University Press.

Ladd, D. R., Faulkner, D., Faulkner, H., and Schepman, A. (1999). 'Constant "segmental anchoring" of F0 movements under changes in speech rate', *Journal of the Acoustical Society of America* 106: 1543-1554.

Ladd, D. R., Mennen, I., and Schepman, A. (2000). 'Phonological conditioning of peak alignment in rising pitch accents in Dutch', *Journal of the Acoustical Society of America* 107: 2685-2696.

Lindblom, B. (1990). 'Explaining phonetic variation: a sketch of the H&H theory', in W. J. Hardcastle and A. Marchal (eds), *Speech Production and Speech Modelling*. Dordrecht: Kluwer Academic Publishers, 403-439.

Malmberg, B. (1943). *Le système consonantique du français moderne: études de phonétique et phonologie. Etudes Romanes de Lund VII.*

Mitleb, F. (1981). *Segmental and Non-segmental Structure in Phonetics: Evidence from Foreign Accent*. PhD diss. Indiana University, Bloomington.

Niebuhr, O. (2003). 'Perceptual study of timing variables in F0 peaks', *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 1225-1228.

Niebuhr, O., and Kohler, K. J. (2004). 'Perception and cognitive processing of tonal alignment in German', in B. Bel and I. Marlien (eds), *Proceedings of the International Symposium on Tonal Aspects of Languages. Emphasis on Tone Languages*. Beijing: The Institute of Linguistics, Chinese Academy of Social Sciences, 155-158.

O'Dell, M., and Port, R. (1983). 'Discrimination of word-final voicing in German', *Journal of the Acoustical Society of America* 73: S31, N17.

Ohala, J. J. (1983). 'Cross-language use of pitch: an ethological view', *Phonetica* 40: 1-18.

_____(1984). 'An ethological perspective on common cross-language utilization of F0 of voice', *Phonetica* 41: 1-16.

Ohala, J. J., and Jaeger, J. J. (eds) (1986). *Experimental Phonology*. Orlando: Academic Press.

Panconcelli-Calzia, G. (1948). *Phonetik als Naturwissenschaft.* Berlin: Wissenschaftliche Editionsgesellschaft.

Paul, H. (1880). *Prinzipien der Sprachgeschichte*. Halle: Niemeyer.

Pierrehumbert, J., and Beckman, M. E. (1988). *Japanese Tone Structure.* Cambridge, MA: MIT Press.

Pierrehumbert, J., and Frisch, S. (1997). 'Synthesizing allophonic glottalization', in J. P. H. van Santen, R. W. Sproat, J. P. Olive, and J. Hirschberg (eds), *Progress in Speech Synthesis*. New York: Springer, 9-26.

Pierrehumbert, J., Beckman, M. E., and Ladd, D. R. (2000). 'Conceptual foundations of phonology as a laboratory science', in N. Burton-Roberts, P. Carr, and G. Docherty (eds), *Phonological Knowledge. Conceptual and Empirical Issues.* Oxford: Oxford University Press, 273-303 .

Port, R., and Crawford, P. (1989). 'Incomplete neutralization and pragmatics in German', *Journal of Phonetics* 17: 257-282.

Port, R., Mitleb, F., and O'Dell, M. (1981). 'Neutralization of obstruent voicing in German is incomplete', *Journal of the Acoustical Society of America* 70: S13, F10.

Port, R., and O'Dell, M. (1985). 'Neutralization of syllable-final voicing in German', *Journal of Phonetics* 13: 455-471.

Prieto, P., d'Imperio, M., and Gili Fivela, B. (forthcoming). 'Pitch accent alignment in Romance: primary and secondary associations with metrical structure', *Language and Speech.*

Rousselot, l'Abbé P. J. (1892). *Les modifications phonétiques du langage, étudiées dans le patois d'une famille de Cellefrouin (Charante)*. Paris: H. Welter.

_____(1897-1901). *Principes de phonétique expérimentale.* Paris: H. Welter.

Scripture, E. W. (1902). *The Elements of Experimental Phonetics.* New York: Scribner's Sons/London: Arnold.

_____(1936). 'Bulletin of the International Society of Experimental Phonetics III', *Archives Néerlandaises de Phonétique Expérimentale* 12: 133-147.

Sievers, E. (1876). *Grundzüge der Lautphysiologie zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Leipzig: Breitkopf and Härtel. (2nd ed. *Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. 1881).

Thorsen, O. (1966). 'Voice assimilation of stop consonants and fricatives in French and its relation to sound duration and intra-oral air-pressure', *Annual Report Institute of Phonetics University of Copenhagen (ARIPUC)* 1: 67-76.

Trubetzkoy, N. S. (1939). *Grundzüge der Phonologie*. (3rd ed. 1962 Göttingen: Vandenhoeck & Ruprecht).