

## PHONETICS - A LANGUAGE SCIENCE IN ITS OWN RIGHT?

K. J. Kohler  
IPDS, Kiel, Germany

### ABSTRACT

This paper starts with some remarks on the history of the ICPhS and argues for a phonetic paradigm in two stages: the heuristics of phonetic phonology, and phonetic explanation. It speaks in favour of phonetics as a language science in its own right on the basis of this paradigm.

### ON THE HISTORY OF THE ICPhS

This scientific meeting is the thirteenth since its inception in Amsterdam in 1932, and it has always been called the International Congress of Phonetic Sciences. At closer inspection, two things are noticed about this name: (a) it refers to a plurality of phonetic sciences and (b) it views this plurality as an open class. In this respect, also because this plurality is meant to include parts of such subjects as psychology, acoustics and linguistics, our Congress differs in a striking way from what is practised in representative disciplines of the Humanities, such as history, or of Science, such as physics. Since a scholarly conference of international dimensions mirrors the theoretical foundations of an academic discipline and the recognition, or absence, of a unified research paradigm [1] constituting a science in its own right, we would have to conclude from the way our Congress has been conceived that the answer to the question of this plenary address is negative.

So I could stop here, and we could all go for a cup of coffee instead. But let us look into this matter more broadly and more deeply and arrive at the proposal of a better-reasoned answer which can at the same time justify - or reject - that we call ourselves phoneticians and that what we do - namely phonetics - is something special. Since the actual state-of-the-art in a subject is always the result of historical incidents and developments it

will help understanding to have a brief look at how and under what auspices this Congress originated.

At the First International Congress of Linguists (The Hague 1928), de Groot proposed that an international periodical of Experimental Linguistics be started, "in order to further the cooperation of Experimental Phonetics, Experimental Psychology and Linguistics, for the study of Language" [2]. In the "Explanatory memorandum" [2] he says: "Instrumental methods are of great importance in nearly every chapter of Linguistic Phonetics, but they need improvement...the phonetician does not always start from a definite linguistic problem; he sometimes even confines the field of Experimental Phonetics to what is of no interest to the linguist at all...his chief interest is often concentrated upon instruments and curves, instead of upon the elements and the functions of speech;..."

The type of experimental phonetics de Groot had in mind was the one practised at his time by such scholars as E. W. Scripture and G. Panconcelli-Calzia. The former saw the 'nature of speech' in measurement-numbers and characterized the phonetic scientist as someone that "might be - and preferably should be - congenitally deaf and totally ignorant of any notions concerning sound and speech." [3, p.135]. The latter explicitly incorporated phonetics into physiology as part of the study of motion, like walking, running, jumping, and therefore regarded phonetics as a natural science, noticing with great satisfaction that the 'philologus auricularius furibundus' of late was getting rare [4, pp. 8,18]. Scripture was present at this congress and succeeded in founding the International Society of Experimental Phonetics

on 11 April, 1928, the day after de Groot made his proposal.

As the president of the International Society of Experimental Phonetics, Scripture planned a first congress of experimental phonetics, which was then held at the Bonn Phonetics Institute in 1930 and organized by P. Menzerath [5]. The second congress of the Society was scheduled in Amsterdam for 1932. However, the Dutch Organizing committee under the chairmanship of the psychologist van Ginneken, and with the phonetician Louise Kaiser as the secretary and the linguist de Groot as a member, decided to invite the "Internationale Arbeitsgemeinschaft für Phonologie" of the Prague Circle, which constituted itself in 1931, and the Amsterdam congress was, therefore, to be "The Second Congress of the International Society of Experimental Phonetics and the First Meeting of the Arbeitsgemeinschaft für Phonologie" as parts of an "International Congress of Phonetic Sciences".

The intentions were clear: the narrow field of the science of experimental phonetics was to be broadened by bringing in the linguistic orientation. This is in keeping with de Groot's summing up [2]: "Phonetics has up to now been too "practical", too didactical; instrumental Phonetics too physiological, too physical, too materialistic; Linguistics too much afraid of instruments." Practical phonetics, of course, referred to the activities of the International Phonetic Association and its prime concern with transcription and pronunciation teaching in foreign languages. So the Dutch organizers had three phonetics branches in mind right from the start: practical phonetics, experimental phonetics and phonology. It was consequently only a small step to broaden the field even further: "After some deliberation and in view of the recent reorganization of the Dutch Society of Phonetics [which in 1931 replaced the Dutch Society of Experimental Phonetics, founded in 1914] we decided that it would be wise to make the sphere of

activity of the congress as extensive as possible and to have phonetic sciences treated in the widest sense." [6] The aim was "that all those who were interested in any aspect of speech sounds should meet and work together" [7].

A circular announcing the congress and its scope was sent out at the end of December 1931, upon which Scripture decided not to hold a Congress of the International Society of Experimental Phonetics. This was the birthday of the International Congress of Phonetic Sciences for short. I think it has now become obvious why the plural was used in the Congress name. At the outset, it refers to no more than a juxtaposition of disciplines, which were still to find the common thread uniting them. This was a task for the future; for the 1932 Congress we are reminded of what Peter Ladefoged said with reference to the IPA: "[it behaves] somewhat like the Church of England - a body whose doctrine is so diffuse that one can hold almost any kind of religious belief and still claim to be a member of it." [8]

### DEVELOPING A PARADIGM OF PHONETICS: FIRST STAGE

#### Integration of phonetics and phonology

We have now explained how the infelicitous name of our Congress originated (which, by the way, also shows a linguistic oddity, no doubt due to an insufficient proficiency of English on the part of the Dutch congress organizers, who translated "wetenschappen" into English, not realising that, contrary to continental usage, English "science" refers to natural science and would normally be in the singular). Other academic disciplines started their congresses after they had reached a common theoretical grounding for all their subsections, expounded in handbooks and expected of anybody wanting to be a member of the same academic circle. In Phonetics it worked just the other way round, and therefore the vital distinction - for the

integrity of a subject - between parts of scholarly activity areas belonging to the same conceptual core, and cooperation of disciplines across their boundaries in questions of mutual concern and interest was blurred.

The question now is whether phonetics has taken this great opportunity of being embedded in an interdisciplinary environment to develop a unifying paradigm that allows a straightforward definition of the subject and its research questions, the setting up of teaching programmes and the publication of comprehensive handbooks of the subject as a whole. The first scholar to reflect thoroughly on the relationship between experimental phonetics and phonology and their integration into what he called the "system of scientific disciplines", was E. Zwirner [9]. His answer was phonometrics [9], which established two essentials: the allocation of measurements to units of language and their statistical evaluation. This view that measurable speech signals are not primarily a physical phenomenon per se but a physical carrier structured for the transmission of meaning in communication has been repeated several times and in various places.

Over the past sixty years the leading centres of phonetic research in the world have established the integration of instrumental and experimental techniques into the context of speech communication. It is a corner stone of modern phonetics that both aspects of human pronunciation, the physical/physiological and the linguistic, are prerequisites of each other. Phonology without a detailed description of the physical manifestation of speech is abstract, and instrumental measurements without their projection onto categories of human communication, linguistic categories among others, are empty and meaningless. Under this view, phonetics includes phonology, albeit a phonology that is at least as closely linked to the laboratory as to the scholar's desk.

### **The linguistic view: linguistic phonology**

So we have certainly advanced since 1932 and created the outlines of a scientific paradigm for phonetics. But in the eyes of linguists, especially of phonologists that are proud of being within the linguistic rather than the phonetic camp, and who advocate - even during keynote addresses at phonetics meetings - that they are not phoneticians the dichotomy between phonetics (conceptualized exclusively as experimental phonetics) and phonology, between Science and the Humanities, persists. Linguists and linguistic phonologists (to coin a term referring to linguists, rather than phoneticians, doing phonology) still regard phonetics as nothing more than the supplier of instrumental data and analyses for the structural slots they have established, i. e. an ancillary appendix to autonomous linguistics, which alone is thought to be capable of giving explanatory accounts of human language. If phonetics is thus devoid of this potential of explaining speech and language phenomena, of the essential ingredient in a scientific discipline, it cannot be a language science in its own right. So, although phonetics has begun to define its own unified basis the attitude of the linguistic world is still that of the thirties.

Even the institution of the Conferences in Laboratory Phonology does not contradict this statement because it is linguistics that is to be taken into the lab to substantiate its categories. The alternative procedure of phonetic measurements obtained and evaluated in the lab being taken into linguistics to confirm, adjust or refute phonological categories by independent assessment is not considered a possibility within this framework.

I would like to buttress this contention with an example from the phonology of German that illustrates the type of epiphenomena that may be created by this 'phonetics-in-phonology' approach. Until Mitleb's thesis of 1981 [10], it was a basic tenet of German phonology that

there is neutralization between word-final voiced and voiceless obstruents. In the interim generative phonology had provided a different account: because of correspondences in morphological paradigms (*Bund* vs. *Bunde*, *bunt* vs. *bunte*) the opposition is postulated at an abstract underlying level for all word positions. Mitleb took this new systematization of the same language phenomena into the lab and tested it with native German speakers who had lived in the US for various lengths of time and who were asked to read word lists containing such unusual items as "Alb" vs. "Alp", but also "weg" vs. "Weck", where there is no morphologically conditioned alternation and "Weck" represents a regionally restricted word. Mitleb being a student of Robert Port's, who in turn learned his phonetics from Leigh Lisker, it is natural that the parameters of 'voicing' he measured were vowel and consonant durations. He found statistically significant differences between the word pairs in the direction expected from the generative description and therefore concluded that the underlying morphophonemic voice distinctions are retained in the production of phonetically voiceless finals through a systematic difference in the length of the preceding vowel: *quod erat demonstrandum*.

However, this finding is the result of a poor methodology of data collection and processing and does not prove anything about the differentiation between these classes of obstruents in the speech of Germans in their native environment, and, of course, says nothing at all about the perceptual relevance of the statistically significant differences as a discriminative function in the communication with a listener. As long as phonology is taken to the lab in this way it will not advance our understanding of how speech communication works, but will simply constitute a self-fulfilling prophecy of autonomous linguistics, which might just as well continue to work with symbolic representations at scholars' desks. That is

what Dinnsen [11] did when he claimed that careful phonetic studies would reveal the non-neutralizing character of perhaps all rules heretofore identified as neutralizations.

But sad to say, even phoneticians fall into this trap set by the way the supremacy of linguistics conceptualizes phonological form and its relation to substance. Francis Nolan [12], at the Second Conference in Laboratory Phonology, having investigated apparent assimilations of final apical to following labial or dorsal stops by electropalatography proposed that differences in lexical phonological form will always result in distinct articulatory gestures, even if overlapped and/or reduced or not discernible in the instrumental record. Here again the questions are as to how good the methodology of data collection was and what these instrumental data can teach us about reduction processes in speech production and their function in communication.

This influence of phonological categorization on the empirical and theoretical work phoneticians do is even more far reaching in the case of Browman and Goldstein's Articulatory Phonology [13]. Their postulates that gestures specified by sets of related tract variables function as primitives of phonological contrast and that gestures are never changed into other gestures, nor added, were undoubtedly triggered by the representation of lexical items with the help of contrastive invariant phonological elements, which are set up in autonomous phonology independently of any function they might have in varying environments of speech communication and which consequently remain invariant. This phonological invariance is extrapolated via the gestural score to the gestures unfolding in articulation.

### **The phonetic view: phonetic phonology**

I have argued against this stand and will do so again in the Symposium on Speaking Styles at this Congress. In

essence my criticism runs as follows. If we, as phoneticians, are interested in gaining insight into how speech communication works, thus transcending the dichotomy of competence and performance, we need to take variants at the phonological level within the same lexical items into account because speakers produce them and listeners successfully decode them. Thus in the German utterance *nun wollen wir mal kucken* ("now let's see") from a dialogue of the Kiel Corpus of Spontaneous Speech [14], displayed in the spectrogram of Figure 1, the phonological citation form representation would be (in IPA transcription) /nu:n 'vɔln 'vi:rə 'ma:l 'kʊkŋ/, but what the speaker pronounced may be symbolized as [nū: ʃñ\* ẽ rja 'kʰokŋ]. This can only give a filtered replica of the articulatory movements that may be deduced from auditory and graphic evaluation of the acoustic record as having taken place. The apical gesture for the second /n/ as well as the two lateral gestures of the phonological representation of canonical word forms have disappeared, whereas the nasalization extends during the whole of what remains of the first three words; the first occurrence of the approximant /v/ has probably left its trace in a labiodental approximation during the vowel sequence of the first two words; the third /n/ is extremely short (37ms) and realised as a tap, and superimposed on it seems to be a labiodental approximation, which may be a continuation of the preceding lip configuration as well as an advance of the same feature in the following /v/, which is not realized as a separate segmental unit; the following vowel is again very short (30ms), and the labiodental gesture continues and tightens to a closure for the following /m/.

I find it impossible to relate this intricate articulatory control to the same invariant gestures as they are to be deduced from the gestural score for the citation forms, in particular the postulate of a gestural reorganization with regard to the

apicals at a higher processing level than the actual articulatory execution seems to be inescapable. Simple temporal sliding and amplitude variation of gestures in the realization of the invariant score cannot explain the empirical facts fully and adequately. On the other hand, separate lexical entries for the words in different communicative environments is out of the question: they are decoded as the same words by the listener (and are, therefore, different from such instances as *zu dem* in *er kam zu dem Schluß, daß...* ("he reached the conclusion that...") versus *zum* in *er kam zum Schluß* ("he came to/at the end"))).

#### Phonology as descriptive heuristics: complementary phonology

In view of these problems with the postulate of invariant phonological units, e. g. phonemes, the question arises as to whether it has contributed anything to the study of language and speech communication. The answer is: "Of course it has!" But the linguistic categories of any phonological model can, at best, only function as heuristic devices, 'As Ifs' in Vaihinger's sense [15], that provide a preprocessing of spoken language data for them to become accessible to further phonetic analysis [16]. Particularly in the case of connected speech, be it read text or spontaneous dialogue, the phonological categorization allows the reduction of a large variability to a small number of entities in canonical word forms that may be listed in a lexicon and to which actually occurring pronunciations are referred. Especially the segmental concept of the phoneme is extremely useful here, e.g. for the labelling of acoustic data bases and for subsequent data retrieval in computer data banks, provided it is integrated with long componential features in a complementary phonology. So in the Kiel Corpus of Spontaneous Speech [14], the utterance of Figure 1 is represented in SAMPA notation as

n u: -MA n+ v- O- l- @- n+ -MA  
v- i:6-6+ m a: l- k -h 'U k @- n-N.

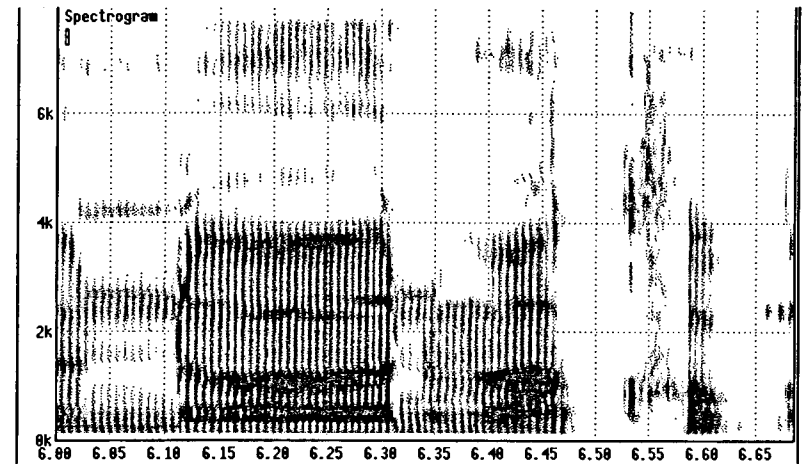


Figure 1. Spectrogram of the German utterance "nun wollen wir mal kucken" from the Kiel Corpus of Spontaneous Speech [14]

For the use of '-' and '-MA' see my contribution to the Symposium on Speaking Styles.

#### PHONETIC EXPLANATION

But none of these phonological devices are explanatory, they are heuristic and descriptive (and, unfortunately, in a large number of purely linguistic phonological studies they are not even that because, as autonomous symbols on paper, they lack the connection with the spoken word). Björn Lindblom has argued on many occasions, e.g. [17], and will certainly develop this point further in the Plenary Symposium on Saturday, that the explanatory questions about speech communication are not answered by phonology as we know it, because it lacks the functional viewpoint with regard to the communicative purpose of speech. It cannot explain why sound systems are the way they are, why speakers change their phonetic output in different situations in the ways they do, and why listeners are still able to decode extremely reduced speech production with great ease. To be able to provide insightful answers to these fundamental questions about speech and language, phonologists would have to step outside their auto-

nomous linguistics field and set up hypotheses that are on the one hand independent of the data to be explained and that are on the other hand related to the biological and social conditions of humans communicating by speech.

In connection with the example of Figure 1 one decisive question is whether any reduction could have taken place at random, or whether the output is structured in highly constrained ways that only allow certain types of deviations from citation form utterances, and this question must be seen under a perspective that goes beyond the individual language, but relates to the physical make-up of the human sound producing system. So the question of language and speech universals is intimately linked to the explanation of individual language data. The specific example comes under three principles: the general instability of apical gestures [18], the greater reduction in word-final than in word-initial position for reasons of word detectability by a hearer, and the greater reduction in non-prominent function words for rhythmic reasons in a stress-timed language like German. Since apical laterals require greater muscular coordination than apical

closures in nasals and plosives, they are more easily dropped than the latter in gestural sequence, when the special conditions for reduction obtain. As the movements of the velum are more sluggish for physiological reasons the nasalization between several nasal-oral-nasal sequences occurs as a matter of course, particularly in fast speech. The superposition of labiodental constriction on tongue body and tip gestures over a relatively large stretch of articulation is made possible by the anatomical and physiological independence of the lower lip and by its slow execution of movements, especially in a repetitive frame /v...v/. So the articulatory manifestations found in this utterance in relation to the canonical forms can be deduced from general principles which would also be applicable to other languages, given the same rhythmic structure and the same tolerance of hearers under social constraints. Historical sound change exemplifies these developments over and over again in the most diverse languages, as John Ohala has pointed out on several occasions, e.g. [19].

The other pertinent explanatory questions related to the utterance in Figure 1 are: "How far do listeners allow degradations of this sort to go before they have to ask for repetition because they do not understand?" and "How do speakers manage to decode such reduced speech correctly and with such ease?" No answers are available as yet. But here is a specific task for phonetics, which falls outside linguistic phonology, which the latter could not handle, and which it would not even be interested in proposing.

#### A PARADIGM OF PHONETICS: SECOND STAGE

So the paradigm of phonetics is taking shape. The integration of phonology and the physics of speech in a phonetic phonology, as expounded above, constitutes the first part of this paradigm: a heuristic framework for phonetic descriptions of languages in all their speech manifestations. Built on this is the second part of

the paradigm: the functional view of speech production and reception - the explanation of the speech communication process between a speaker and a listener (and we may add, the acquisition of language and speech) with reference to the physics, biology and social environment of homo loquens. No other discipline has or wants to have such a paradigm. Linguistics is content within its autonomous framework detached from the purpose it may be put to in communication; acoustics and engineering (except for the engineers that have adopted, or in the case of the colleagues at KTH even assisted in creating, the phonetic paradigm) are only interested in the physical perspective, as is illustrated, for example, by the way they deal with automatic speech recognition or with building block synthesis.

Picking up the theme of Francis Nolan's paper on "Phonetics in the next ten years" at the last Congress [20], I would now venture to say that the coming years will see a consolidation of this paradigm of modern phonetics as a unitary discipline of the spoken medium of language, an essential interface between the pure and simple signal approach of physics and engineering and the symbolic orientation of semantics, syntax and linguistic phonology in linguistics. Phonetics will thus occupy a key position in enquiries into the functioning of speech communication at the levels of pure research as well as application. Of course, there must and always will be interdisciplinary cooperation with neighbouring fields that have different paradigms, but can contribute special expertise which the phonetician does not have, e.g. acoustics, physiology, psychology, linguistics.

This paradigm also necessitates the training of phonetics students in symbol as well as signal aspects of speech and language, including analytic listening and transcription techniques, speech signal processing and experimental methods. A common core curriculum will be developed and a "Handbook of Phonetic

Science" will be written in accordance with the paradigm. There are already initiatives for a phonetics curriculum at the European level of the ERASMUS programme of Phonetics and Speech Communication, although its compilation of subject areas is still too encyclopedic and juxtapositional with not enough reference to the phonetic paradigm.

#### CONCLUSION

I can now go back to the title of this talk. Yes, phonetics, in my view, is a language science in its own right by virtue of its subject matter, and it is well on its way towards asserting itself as such. There is still a good deal of hard work ahead of us. Let's begin with an evolution of our historical traditions and drop just one letter and two phonemes, at the end of the Congress name!

#### REFERENCES

- [1] Kuhn, T.S. (1970), *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.
- [2] de Groot, A. W. (1928), "Proposition 2 - Explanatory memorandum", *Actes du Premier Congrès International de Linguistes*, pp. 6-9, Leiden: Sijthoff's Uitgeversmaatschappij.
- [3] Scripture, E.W. (1935), "Bulletin of the International Society of Experimental Phonetics III", *Arch Neerl Phon Exp*, vol. 11, pp. 133-147.
- [4] Panconcelli-Calzia, G. (1948), *Phonetik als Naturwissenschaft*, Berlin: Wissenschaftliche Editionsgemeinschaft.
- [5] Menzerath, P. (1930), *Bericht über die I. Tagung der Internationalen Gesellschaft für experimentelle Phonetik*, Bonn: Universitäts-Buchdruckerei Gebr. Scheur.
- [6] (1932), *Proceedings of the International Congress of Phonetic Sciences*, Amsterdam.
- [7] Fischer-Jørgensen, E. (1984), "Some aspects of the 'Phonetic Sciences', past and present", in M.P.R. van den Broecke, A. Cohen (eds.), *Proc Xth Inter Cong Phon Sc*, pp. 3-11, Dordrecht: Foris Publications.

- [8] Ladefoged, P. (1990), "Some reflections on the IPA", *Journal of Phonetics*, vol. 18, pp. 335-346.
- [9] Zwirner, E., Zwirner, K. (1936), *Grundfragen der Phonetik*, 2nd ed., Basel: Karger, 1966.
- [10] Mitleb, F. (1981), *Segmental and non-segmental structure in phonetics: evidence from foreign accent*, PhD. diss. Indiana University, Bloomington.
- [11] Dinnsen, D.A. (1983), *On the characterization of phonological neutralization*, Bloomington: Indiana University Linguistics Club.
- [12] Nolan, F. (1992), "The descriptive role of segments: evidence from assimilation", in J. Docherty, B. Ladd (eds.), *Papers in Laboratory Phonology II*, pp. 261-280, Cambridge: CUP.
- [13] Browman, C.P., Goldstein, L. (1992), "Articulatory phonology: an overview", *Phonetica*, vol. 49, pp. 155-180.
- [14] IPDS (1995), *CD-ROM#2: The Kiel Corpus of Spontaneous Speech*, vol. I, Kiel: IPDS.
- [15] Vaehinger, H. (1920), *Die Philosophie des Als Ob*, Leipzig: Meiner. (Transl. Ogden, C.K. (1965), *The Philosophy of 'As If'*, London: Routledge & Kegan Paul.)
- [16] Kohler, K.J. (1991), "The phonetics/phonology issue in the study of articulatory reduction", *Phonetica*, vol. 48, pp. 180-192.
- [17] Lindblom, B. (1980), "The goal of phonetics, its unification and application", *Phonetica*, vol. 37, pp. 7-26.
- [18] Kohler, K.J. (1976), "Die Instabilität wortfinaler Alveolarplosive im Deutschen - eine elektropalatographische Untersuchung", *Phonetica*, vol. 33, pp. 1-30.
- [19] Ohala, J. (1983), "The origin of sound patterns in vocal tract constraints", in P.F. MacNeilage (ed.), *The Production of Speech*, New York/Heidelberg/Berlin: Springer.
- [20] Nolan, F. (1991), "Phonetics in the next ten years", *Proc XIIth Inter Cong Phon Sci*, vol. 1, pp. 125-129, Aix-en-Provence: Université de Provence.