Review Article

From Phonetics to Phonology and Back Again

Ladd, D.R.: Phonetics in Phonology; in Goldsmith, Riggle, Yu, The Handbook of Phonological Theory; 2nd ed., pp. 348–373 (Wiley-Blackwell, Chichester 2011)

Klaus J. Kohler

University of Kiel, Kiel, Germany

Defining the Scope

The new edition of Goldsmith [1995] has a very different content structure, for example, it no longer contains chapters on 'Experimental Phonology' (Ohala) or on 'The Internal Organization of Speech Sounds' (Clements and Hume), and it advocates a different perspective on the relationship between phonology and phonetics, summarized as follows in the 'Preface':

Comparing the present handbook to the one that was produced in 1995, we seem to find, too, that the field has expanded: it now includes a good deal more content and emphasis on phonetics, on variation, and on computational approaches. In reality, the growth is more a matter of perspective than anything else: studies on phonetics, variation, and computation that were of interest to phonologists have existed for a long time, but the perception is now much stronger that this work is not outside the field of phonology (though of interest to some psychologists), as it is a real and integral part of the field itself (*Hb*, pp. ix–x).

The title of Robert Ladd's chapter 11, 'Phonetics in Phonology', bears witness to the changed outlook, and the editors introduce it like this:

D. Robert Ladd's chapter on the role of phonetics in phonology is a good example of how the thematic questions at the center of phonological discussions have evolved over the past 15 years. The time depth of his discussion, involving scholars working over almost all of the twentieth century, is considerably deeper than that found in many of the chapters of the 1995 volume, and Ladd explicitly draws together the views that Trubetzkoy developed in the 1930s with those at the heart of classical generative phonology and those that scholars today are developing, often under the influence of far richer computational resources than was imaginable even 25 years ago. The easy assumption that phonetic reality can be modeled with a well-designed symbolic representation, such as that produced by the International Phonetic Association, have been widely challenged, and Ladd asks what alternative accounts are available to us now for characterizing the nature of phonetic reality (*Hb*, pp. xi–xii).

This chapter is thus given a central place in the development of phonological theory as documented in the second edition of the *Handbook of Phonological Theory*, especially as regards a new view of the relationship between phonology and phonetics. It is, therefore, justified to review it on its own in a phonetics journal, and to examine whether speech science has moved in the right direction since Trubetzkoy divided the

Prof. Dr. Klaus J. Kohler
Holm 4
DE–24113 Molfsee (Germany)
E-Mail kjk@ipds.uni-kiel.de

field into two equipollent disciplines, phonetics and phonology. In particular, the subjugation of phonetics to an abstract edifice of formal structures, as Ladd advocates it with 'Phonetics in Phonology', will be under scrutiny.

For L. the dichotomy of phonetics and phonology builds on 'The dual nature of speech sounds – as physical events and as elements of a symbolic system' (*Hb*, p. 348), separately reflecting the cognitive representation of discrete categories in a language and their continuous parametric physical manifestations in speech. In the vein of generative phonology [Chomsky and Halle, 1968], L. aims at an overall formal model of language and speech that explicates the transformation of underlying symbolic categories of *langue* into physical time courses of *parole*, giving the phonological level logical precedence over phonetic manifestation. But he deviates from generative phonology, where the output of the phonological component of a language is a systematic phonetic representation of speech sounds (phones) from a universal repertoire, which, in turn, is the input to a phonetic implementation system, outside *langue*, to generate the final physical output. Instead, L. wants to incorporate the physical phonetic level into the phonological component and argues against the need for, and the usefulness of, a systematic phonetic representation in his encompassing model of the sound of language and speech. His argumentation proceeds along a historical survey of

– Trubetzkoy's [1939] two sound levels – sound in speech acts (*Sprechaktlautlehre*) and sound in language systems (*Sprachgebildelautlehre*);
– the concept of speech sounds in phonetic transcription, particularly the International Phonetic Alphabet (IPA) – the alphabet of the International Phonetic Association;
– the concepts of systematic phonetics, of the phone, and of distinctive features, and
– the interface between the levels of discrete phonology and continuous phonetics.

L.'s contribution raises three questions, which this review tries to answer: (1) Does the historical survey provide a faithful account of phonetic concepts in relation to phonology? (2) What type of formal model does L. propose for phonology? (3) Is 'Phonetics in Phonology' a convincing solution to the phonology-phonetics dichotomy? Is it appropriate to absorb *phonetics* into *phonology*, considering that an adequate evaluation of sound patterns of *homo loquens* needs to take the primacy of relationships between Speaker, Listener and Communicative Situation into account, over and above, and independently from, the sum of the formal relations?

Phonetic Concepts in Relation to Phonology

Trubetzkoy's Two Sound Levels

L.'s discussion takes its point of departure from Trubetzkoy's *Grundzüge der Phonologie* and refers to the second German edition of 1958, the English translation [Baltaxe, 1969], and the French translation [Cantineau, 1949]. Quotes and page numbers are generally from the English translation. The lack of page references to the German original is a regrettable omission in a scholarly treatise which puts Trubetzkoy's ideas and terminology of sound in phonetics and phonology in focus for a handbook presentation of a phonological framework. In the following discussion, I shall refer to the German original with page numbers prefixed by *Gr*, and with my own English translations. Page references to L.'s statements are prefixed by *Hb*.

With regard to the division of *Lautlehre* into the two strictly separated levels of *Sprechaktlautlehre* and *Sprachgebildelautlehre* (*Gr*, pp. 5ff.), L. attempts his own translation as 'the study of the sounds of [Saussurean] *parole*' and 'the study of the sounds of *langue*' (*Hb*, pp. 348f.). The use of the plural 'sounds' is wrong and misleading, because Trubetzkoy puts *Lautlehre* beside *Bedeutungslehre*, 'study of sound' versus 'study of meaning', as fields in linguistics, and therefore *Sprechaktlautlehre* is 'the study of sound in speech acts' (= phonetics), *Sprachgebildelautlehre* 'the study of sound in language systems' (= phonology). This means that for Trubetzkoy both fields deal with 'sound'.

'Sound in speech acts' comprises the 'sound stream of the concrete speech act… an uninterrupted, seemingly unordered sequence of intertwined sound movements' (*«Der Lautstrom des konkreten Sprechaktes ist eine ununterbrochene, scheinbar ungeordnete Aufeinanderfolge ineinander übergleitender Schallbewegungen»*, *Gr*, p. 6), as well as 'typical articulation and sound elements excerpted from the sound and articulatory continuum' (*«aus dem lautlichen und artikulatorischen Kontinuum herausgegriffene[n] typische[n] Artikulations- und Schallgebilde»*, *Gr*, p. 17). These excerpts are speech sounds, which include all sound information of the speech act at that point.

'Sound in language systems', on the other hand, 'only takes into account what fulfils a specific function in the language system' (*«nur dasjenige ins Auge fassen, was eine bestimmte Funktion im Sprachgebilde erfüllt»*, *Gr*, p. 14), i.e. phonemes, comprising 'the sum of phonologically relevant properties of a sound element' (*«die Gesamtheit der phonologisch relevanten Eigenschaften eines Lautgebildes»*, *Gr*, p. 35). Thus, there are no sounds in 'the study of language systems':

> Every concrete sound, produced and perceived in the speech act contains a lot of other, phonologically irrelevant properties, besides the phonologically relevant ones. Therefore none of these sounds can simply be equated with a phoneme… The phonemes are realized by speech sounds, of which every speech act is composed. These speech sounds are never the phonemes themselves, since a phoneme must not contain phonologically irrelevant traits, which is inescapable for an actually generated speech sound. (*Jeder von den konkreten im Sprechakt erzeugten und wahrgenommenen Lauten enthält außer den phonologisch relevanten noch viele andere, phonologisch irrelevante Eigenschaften. Daher kann keiner von diesen Lauten kurzweg als Phonem betrachtet werden. … Die Phoneme werden durch Sprachlaute (genauer Sprechlaute, Redelaute) realisiert, aus denen jeder Sprechakt besteht. Diese Sprachlaute sind niemals die Phoneme selbst, weil ja ein Phonem keine phonologisch irrelevanten Züge enthalten darf, was für einen tatsächlich erzeugten Sprachlaut unvermeidlich ist, Gr*, pp. 35f.)

Here Baltaxe's [1969, p. 4] English translation, although clumsier, is a more faithful rendering of the German text than L.'s translation: 'the study of sound pertaining to the act of speech' and 'the study of sound pertaining to the system of language'. The French translation is equally faulty: 'la science des sons de la parole' and 'la science des sons de la langue' [Cantineau, 1949, p. 3]. This difference between 'sound' in both fields and '(speech) sounds' in phonetics is crucial to understanding Trubetzkoy's partitioning of the 'study of sound' into two fields, according to their association either with the physical aspects of sound in speech acts, by the a-semantic methods of the natural sciences (especially the practice of the experimental phoneticians of the late 19th and the early 20th century), or with the functions of sound in a particular language, by the semantic methods of linguistics and the social sciences. In spite of their fundamental difference, due to their different perspectives and methods, links between the two fields are 'inevitable and absolutely necessary' (*«Somit ist ein gewisser Kontakt zwischen der Phonologie und*

Kohler

der Phonetik trotz ihrer grundsätzlichen Unabhängigkeit unvermeidlich und unbedingt notwendig», *Gr*, p. 17).

'At the outset, every phonological description consists in the discovery of the existing sound contrasts that differentiate meaning in the particular language. For this, the phonetic record of the particular language has to be the point of departure and provide the material.' (*«Der Anfang jeder phonologischen Beschreibung besteht in der Aufdeckung der in der betreffenden Sprache bestehenden bedeutungsdifferenzierenden Schallgegensätze. Dabei muß die phonetische Aufnahme der betreffenden Sprache als Ausgangsunkt und als Material genommen werden»*, *Gr*, p. 17).

However, at this point the concept of 'abstractive relevance' (*«abstraktive Relevanz»*, *Gr*, p. 49) intervenes, which Trubetzkoy took over from Karl Bühler [1931, p. 38], and which refers to abstracting the phonologically relevant sound properties from the totality of sound. This then leads to the 'phonological essence of a phoneme' (*«phonologischer Gehalt eines Phonems»*), 'the sum of all phonologically relevant properties of a phoneme, i.e. of all those properties, which all variants of this phoneme have in common, and which differentiate it from all other … phonemes of the same language… The phonological essence of the German *k* phoneme can only be labelled as follows: "tense non-nasal dorsal stop".' (*«Unter phonologischem Gehalt verstehen wir den Inbegriff aller phonologisch relevanten Eigenschaften eines Phonems, d. i. jener Eigenschaften, die allen Varianten dieses Phonems gemeinsam sind und es von allen anderen … Phonemen derselben Sprache unterscheiden… Der phonologische Gehalt des deutschen k-Phonems läßt sich nur so formulieren: ‹gespannter nichtnasaler dorsaler Verschlußlaut›»*, *Gr*, p. 59.) In general, these properties are aspects of sound. But since the definition of the phonological essence of a phoneme depends on the position the phoneme occupies in the phoneme system , i.e. on the other phonemes it contrasts with, 'it can *sometimes* [my emphasis] receive a completely negative definition' (*«Daher kann ein Phonem manchmal eine rein negative Definition erhalten»*, *Gr*, p. 60), such as 'non-lateral liquid' for all the facultative and combinatory variants of the German *r* phoneme. But the phonetic essence is usually linked to sound.

This is another point where L. misrepresents Trubetzkoy's phonology-phonetics relationship, as he states (*Hb*, p. 355): 'Trubetzkoy starts out by describing phonology in purely abstract terms ("The signifier of the system of language [i.e. of langue] consists of a number of elements [viz., phonemes – D.R.L.], *whose essential function it is to distinguish themselves from each other*." (Baltaxe 1969: 10, emphasis added)).' But L. leaves out the important subsequent sentences:

Every word needs to be differentiated from all other words of the same language system *by something* [my emphasis]. … The phonology has to analyse which differences in sound are associated with differences in meaning in the particular language, how these differentiating elements (or marks) are related to each other and according to what rules they may be combined to words (or sentences). (*«Das Bezeichnende des Sprachgebildes besteht aus einer Anzahl von Elementen, deren Wesen darin liegt, daß sie sich voneinander unterscheiden. Jedes Wort muß sich von allen übrigen Wörtern desselben Sprachgebildes durch etwas unterscheiden. … Die Phonologie hat zu untersuchen, welche Lautunterschiede in der betreffenden Sprache mit Bedeutungsunterschieden verbunden sind, wie sich die Unterscheidungselemente (oder Male) zueinander verhalten und nach welchen Regeln sie miteinander zu Wörtern (bezw. Sätzen) kombiniert werden dürfen»*, *Gr*, p. 14).

This 'something', these 'marks' are aspects of sound, not abstract symbols. L. confuses 'abstractive relevance' with 'abstract representation'. And it is not correct either – what L. maintains (*Hb*, p. 355) – that for Trubetzkoy reference 'to concrete phonetic properties of phones', instead of 'to abstract dimensions', is 'merely an

expedient codification of the dimensions of the IPA chart'. For Trubetzkoy these differentiating marks in phonology are empirically rooted in differences of sound, and they do not contradict his theoretical and methodological separation of the two levels in the study of sound. So, L. accuses him wrongly of not taking his 'radical talk of the distinction between phonology and phonetics seriously' (*Hb*, p. 368), and of treating the rigid separation he posited as an illusion (*Hb*, p. 358).

How sound properties in phonological relations determine the phonological essence of a phoneme can be seen in the concept of the 'archiphoneme' (*Gr*, p. 71). When a phonological opposition is neutralized in a certain word position or context, such as lenis and fortis obstruents word-finally in German, the sound property that distinguishes the two members of the opposition outside the neutralizing environment can no longer be part of the phonological essence of the 'archiphoneme' under neutralization, which is, in turn, represented by the speech sound occurring in this neutralizing position. This is very different from postulating an underlying *abstract* opposition in all contexts from which the concrete neutralized member is derived, e.g. by rule, as in generative phonology. L. does not discuss the archiphoneme concept and its implicit polysystemic approach of phonological analysis, which was further developed by the Firthian School of Prosodic Analysis [Ogden and Local, 1994; Hawkins and Smith, 2001], and needs to be integrated into a modern theory of *phonetics in phonology.*

The lack of understanding that Trubetzkoy does not postulate 'sounds of *langue*', but phonemes containing a phonological essence on the basis of abstractive relevance of sound, finally leads L. to fanciful speculation about abstract relations:

> Considering the importance that Trubetzkoy attached to the phonology-phonetics distinction, the persistence of traditional phonetic dimensions in phonology is striking. One could perfectly well imagine a description of the distinctive oppositions in a given language that makes no reference to phonetics and really does work with the idea of abstract distinctness or 'mere otherness'. Standard names for the four tonemes of Mandarin Chinese are essentially of this sort: the long-standing Western practice of using the numbers 1 to 4 obviously makes no reference to the pitch contours by which the abstract tonemes are phonetically manifested. … Indeed, this might seem to be a good way of pursuing Trubetzkoy's professed goal of categorizing 'the sounds of *langue*': such names or numbers are shorthand ways of referring to abstract phonological elements that are functionally equivalent across the lexicon irrespective of phonetic realization' (*Hb*, p. 357).

This completely misses the theoretical and descriptive orientation of Trubetzkoy's phonological work on a huge number of languages, which he draws on in phonetic terms in his *Grundzüge*.

IPA Transcription and Systematic Phonetics

L. makes several statements about the IPA of the International Phonetic Association which are incorrect or inadequate. A minor formal point concerns the reference to 'successive editions of the IPA *Handbook*' and 'the handbook's most recent edition [IPA, 1999]' (*Hb*, p. 349) – there has only been one edition of the comprehensive *Handbook*, which was preceded by a small pamphlet *The Principles of the International Phonetic Association*, last revised in 1949 [IPA, 1949]. Fundamental problems in L.'s argumentation arise in his discussion of the concepts of *systematic phonetics* and of the *phone*, which he relates to the 'IPA enterprise' (pp. 350, 365). This phrasing, put in the context of the final statement of the chapter, gets a negative connotation of something unsuccessful and not to be taken seriously as a 'theoretical basis for describing utterance phonetics' (*Hb*, p. 365): 'I don't believe that we will make much progress in phonology

until we stop trying to ground our theories in the systematic phonetic representation of individual acts of speech' (p. 371). This is turned into a derisive attitude towards IPA phonetics in a reference to the 'soul-searching [that is] still go[ing] on among self-identified phoneticians thoroughly familiar with the continuous parametric nature of speech' (*Hb*, fn. 4, p. 372).

According to L., systematic phonetics depends on two key premises, the *segmental idealization* and the *universal categorization* assumption. He defines the former assumption as 'Speech (NB not language) can appropriately be idealized as a string of ordered discrete sound segments of unspecified duration' (*Hb*, p. 349), with reference to IPA [1999, p. 6] (not 5, as given in the quote): 'Phonetic analysis is based on the crucial premise that it is possible to describe speech in terms of a sequence of segments'. The second assumption is given as 'There is a closed universal inventory of possible segment types' (*Hb*, p. 350), with reference to IPA [1999, p. 159]: 'The IPA is intended to be a set of symbols for representing all the possible sounds of the world's languages'.[1] The two assumptions are referred to two different levels within IPA-type phonetics: (1) segmental analysis of speech and classification of the segments according to a limited number of categories, (usually) determined by the speech producing apparatus of *homo loquens*, and (2) the graphic representation of these classified segments by a systematic phonetic alphabet on a Roman basis. Thus, (1) contains both assumptions, and it is older than IPA categorization, e.g. Bell [1867], Brücke [1856], Sievers [1876], Sweet [1877], who laid the foundation for the practical application of IPA sound classification and transcription in foreign language teaching and learning.

The origins of sound classification go back several thousand years to the invention of alphabetic writing in the Semitic language family, where words are grouped to semantic fields in the lexicon and represented by phonetic sameness of three consonants in a particular sequential order, such as **ktb** for the field of WRITING [Mitchell, 1962]:

kˈa**t**a**b**	he wrote	**k**ˈa**t**a**b**a	clerks	ma**kt**ˈuu**b**	written
yˈi**kt**i**b**	he writes, will write	**k**i**t**ˈaa**b**	book	mˈa**kt**a**b**	office, desk
kˈaa**t**i**b**	clerk	**k**ˈu**t**u**b**	books	ma**kt**ˈa**b**a	library

L. will no doubt say that this is an obvious case of phonology, more specifically of phonemics, and that it does not point to a theoretical foundation of systematic phonetics. But what the inventors of the first alphabetic writing had to achieve was to extract certain segmental units from the stream of sound and establish their sameness, with reference to consonantal places and strictures, *not* their otherness, and that is a fundamentally different principle from the 20th century phonemic one; as a matter of fact, it is a basic concept of systematic phonetics. Of course, this classification and sound-to-symbol conversion needs to abstract from time, apart from serial ordering, and ignore properties that do not identify the three-consonant semantic root, such as those caused by gender, expressiveness, attitudes, context. But this systematic phonetic specification serves a useful function, namely a systematic graphic representation of speech. The Devanagari script, which built on Semitic writing, although it no longer

---

[1] It is strange that the second assumption is linked to a quotation from Appendix 1 of the *Handbook*, which contains archival material about the 1989 Kiel Convention of the International Phonetic Association. Since the quote refers to the use of the IPA to transcribe any language it should have been taken from the main body of the *Handbook* that deals with this goal: 'it is desirable to have a consistent way of representing the sounds of language in written form' (p. 3). This refers to the language faculty of *homo loquens*.

had the semantic backing, also focussed on consonantal sounds in the sound stream and attached vocalic modifications of syllables to the basic consonant symbols. And the Sanskrit grammarians, especially Paṇini (about 300 BC), gave very detailed segmental phonetic classifications [Whitney, 1896].

The IPA adopted the general segmental phonetic systematization and expanded it to be applied to any language. It is either a descriptive tool at various levels of delicacy, captured by the terms *broad* and *narrow* transcriptions, with degrees of narrowness [Abercrombie, 1964]; or it is a discovery tool for reducing unwritten languages to writing in field work, thus assisting the development of national orthographies, as happened, for example, in West Africa. The discovery tool has to be an impressionistic transcription with close reference to phonetic categories, but as fieldwork progresses, considerations of sound contrasts quickly intervene, eventually leading to a phonemic-type transcription, which can be modified again in a narrow transcription by adding more and more non-contrastive phonetic detail in the symbolic representation. This is the progression from phonetics to phonology, which Trubetzkoy referred to when he said that the phonetic record has to be the point of departure and provide the material to ascertain the existing sound contrasts. This is also what Pike [1947, p. 57a] meant when he said: 'Phonetics gathers raw material, phonemics cooks it'.

The phonetics used in the field is first and foremost systematic phonetics, applied by a phonetician who has been thoroughly trained in producing, perceiving and transcribing the general phonetic sound categories. 'The Association strongly recommends that anyone intending to use the symbols should receive training in order to learn how to produce and recognize the corresponding sounds with a reasonable degree of accuracy' [IPA, 1999, Appendix 1, Resolutions of the 1989 Kiel Convention, p. 160]. Moreover, any IPA transcription, whether impressionistic or broad/narrow systematic, consists of two parts, the string of symbols and a set of conventions, the more detailed the symbolization the fewer the conventions. Some phoneticians, for example in Firthian Prosodic Analysis [Ogden and Local, 1994], make extensive use of additional comments to their segmental transcriptions, capturing such properties as long range articulatory prosodies of palatalization, velarization, labialization, nasalization.

Admittedly, not every person professing to be a phonetician has the same level of proficiency in systematic phonetic analysis; Daniel Jones, Ian Catford, Kenneth Pike, Peter Ladefoged were exceptional. Very few of modern phonologists would be able to produce a transcription that goes beyond a phonemic one, because they have not had the training through which they acquire a fairly objective reference system of general phonetic auditory and articulatory categories to which they can relate the sounds they encounter in the speech of individual languages. But the conclusion cannot be L.'s [*Hb*, fn. 3, p. 372]:

It is probably not quite accurate to say that phone abstractions are based on the personal equipment of individual phoneticians; it would be better to describe them as the collective effect of the personal equipment of a group of scholars who were all literate in alphabetic writing systems and all spoke more than one European language. Alphabetic literacy inclined them toward the segmental idealization; familiarity with several languages that used the same alphabet inclined them to identify cross-linguistic categories of sound like [b] and [p] and to focus their attention on specific phonetic details (like the difference between [p] and [ph]) that were salient in the comparative description of the European languages.

Influence from the analysts' native and acquired languages can certainly never be excluded completely, but to put all IPA-type phoneticians in the straightjacket of European languages with alphabetic writing, from which there is no escape, ignores at

least 150 years of successful systematic phonetic analysis of speech and languages, and the independent evidence of segmental analysis and classification in the centuries beforehand. This is the modern phonologist's reversal of the route from phonetics to phonology.

The Phone

From the two assumptions of systematic phonetics, L. derives its key theoretical construct of the *phone*.

The phone has been part of the IPA enterprise from the very beginning, … the first version of the IPA alphabet was published in 1888, along with a set of principles on which it was based. The first of these principles [IPA, 1999, p. 198; Appendix 4, The History of the International Phonetic Association] was:

'There should be a separate sign for each distinctive sound; that is, for each sound which, being used instead of another, in the same language, can change the meaning of a word.'

In modern terms, this very clearly states that IPA transcription is intended as a phonemic transcription, and sound is clearly being used to mean 'phoneme'. However, the seeds of theoretical confusion were sown immediately, in the second principle:

'When any sound is found in several languages, the same sign should be used in all. This applies also to very similar shades of sound.' [ibid.].

… The notion of sound in the first principle is language-specific; the notion of sound in the second implies a language-independent categorisation. This second sense of sound is what came to be known as the phone (*Hb*, p. 350).

Contrary to the impression L. gives in the above quote, the term *phone* is not used in the *IPA Principles* or the *IPA Handbook*, nor by Daniel Jones, nor by Trubetzkoy. There the general term is *(speech) sound*. Bloch [1948] has *segment* instead. But the term does occur in some early American publications on phonology, e.g. Hocket [1942, 6.1], 'A phoneme is a class of phones …', or Bloch [1950, 0.2]:

An ALLOPHONE is a subclass of a phoneme, composed of phones that are the same. In this paper 'phone' will be used ambiguously in two senses: strictly to designate a unique event; and loosely, to designate a class of phones that are the same. In the latter meaning, the term 'phone' is a shorter equivalent of 'allophone'.

Finally, Pike [1943, p. 115] defines several types of phone in purely phonetic terms, in keeping with his strict separation of categories of *Phonetics* [Pike, 1943] and *Phonemics* [Pike, 1947]. Two segments are the same *phonetic unit* or *phone* when they are produced by the same articulatory method. and are acoustically the same. 'An *instrumental phone* is a phone identified or identifiable by some instrumental means.' 'A *real phone* is one which the average normal ear, after training, elimination of phonemic prejudice, and so on, would identify, or be physiologically capable of identifying.' 'A *perceptual phone* is one which a particular ear at a particular time believes it identifies.' Joos [1948] also uses *phone* in the phonetic sense with specific reference to Pike [1943].

L.'s dubious argumentation against IPA systematic phonetics as a 'theoretical basis for describing utterance phonetics' is thus built on the ill-defined term *phone*, which oscillates between phonetics and phonology in American usage and does not belong to the phonetic framework of the IPA. It is furthermore highly unfortunate that L. quotes two IPA principles in their wording of 1888 but with reference to the historical Appendix of the 1999 *Handbook*, although the *Handbook* itself no longer lists these principles, and worse still, that he then quotes part of their rephrased versions from the

1949 *Principles*, leaving out the important changes. The two principles cited above do not deal with phonetic analysis but with alphabetic symbolization. It was the decided policy to codify an alphabet for phonetic transcriptions that can be printed with as little delay and cost as possible, and can be read with ease. In those early days, for every new character, including combinations of a main symbol with diacritics, a new metal sort had to be produced costing time and money. Therefore, the basis of the IPA was to be Roman, diacritics were to be avoided, and the same symbol was to be used whenever possible. Thus, similar sounds across languages were to be given the same symbol, but sounds that distinguished the meanings of words had to have different symbols. Such typographically simple transcriptions also assist easy reading.

In the 1949 revision of the two principles, the typographic restrictions were speci-fied, in particular, the second principle now contains the following addition: 'Separate letters or diacritical marks may, however, be used to distinguish [acoustically close sounds] in "narrow" transcriptions or in scientific investigations' [IPA, 1949, p. 1]. And this distinction between 'broad' and 'narrow' transcriptions was given a whole section in the 1999 *IPA Handbook* (pp. 28–30). L. should have quoted the following key sen-tences: 'If a transcription is made in circumstances where nothing can be assumed about the phonological system, it is necessary to include all phonetic details because it is not clear which phonetic properties will turn out to be important. The transcription would be made taking into account only the phonetic properties of the speech. This type of narrow transcription, as might be made in the first stages of fieldwork … is sometimes called an impressionistic transcription or a general phonetic transcription' (pp. 28f.). This IPA principle coincides with the one advocated by Trubetzkoy, and L.'s whole argument would have become superfluous had he taken these issues into account.

L. then moves on to the central role that systematic phonetics came to play in theo-retical phonology during the 1930s and 1940s, when 'the phone idealization became firmly embedded in linguistic discussions of the phoneme on both sides of the Atlantic' (*Hb*, p. 352). Talking about the realization of voiceless stops initially and after /s/, he refers, on the one hand, to the instrumental measure voice onset time (VOT) of 50–70 ms in the former position, and of 0–20 ms in the latter, and, on the other hand, to the phone representation with IPA transcriptions [pʰ] and [p]. He considers the phones 'a considerable abstraction away from observations about VOT', although accounts of the phoneme treat them as 'raw data' and as providing 'a faithful representation of what a speaker really produces … Rather than recognize [pʰ] and [p] as abstractions based … on the perceptual equipment of the transcriber, classical phoneme theory took them as categories of phonetic description, identifiable in a language-independent way' (*Hb*, p. 352).

Inverting this statement leads to the conclusion that L. takes the non-phone-related instrumental VOT measure as the raw data and a faithful representation of what the speaker really produces. He obviously does not realize that VOT duration is also a con-siderable, albeit different, abstraction from the time courses of articulatory parameters. For a start, it is phone-like segmental *par excellence*, lumping the dynamics from stop release to onset of voice phonation together in a building block with a specified dura-tion, which in IPA terminology is 'no, weak or strong aspiration', [p], [pʰ] or [ph], the only difference being the absence of a time measure. And, strictly speaking, the concept only applies to pre-sonorant position; utterance-finally, e.g. in the English citation form 'top' [pʰ], there is no voice onset following the plosive release. But from the point of view of speech production, this building block is dynamically structured as a sequence

of release burst, local friction and cavity friction, with different temporal extensions and acoustic energies of the three parts at labial, dental, alveolar, palatal, velar, uvular places of articulation. At the alveolar place, the local friction may be caused in a slit-like stricture, leading to a [θ]-type fricative, but at the alveolar ridge, rather than the teeth, as at the end of a word like 'night' in Irish English, or it may be caused by a groove-like stricture, assuming acoustic properties of [s], which are particularly strong in Danish [tˢ], and which resulted in the Old High German sound change [t] > [ts], [s]. To find out what the speaker really does *VOT* performs no better than *aspiration*.

But the crucial question is what we want to achieve with the statements we make about speech in a language, at various degrees of delicacy. A well-trained phonetician can easily perform and recognize the series of phonetic categories of occlusion + release in plosive articulations, ranging from voiced phonation during the occlusion in [b] to complete voicelessness during the occlusion followed by a release burst in [p], to subsequent short and weak aspiration in [pʰ], to subsequent long and strong aspiration in [ph]. And these phonetic points can easily be projected on a scale from negative to long positive VOT. But the phonetician can also distinguish between a completely voiceless unaspirated stop with either a weak, lenis release in [b̥] or a strong, fortis release in [p], which are outside the VOT scale.

In the comparison of English and French plosives, for example for language teaching, the phonological characterization as [±voice] is completely misleading since a contrast of phonetic voicing only applies to French, in English the contrast is one of aspiration prevocalically, and one of vowel occlusion timing postvocalically. Furthermore, the completely voiceless [b̥], frequently occurring utterance-initially in English, differs from unaspirated French [p], as lenis versus fortis (voiceless unaspirated). It has been shown in Kohler [1981] that when voicing is successively removed from the occlusion of a naturally produced French [b], French-speaking listeners in a contextualized perception test «Quelle est la lettre suivante? bé.», become uncertain as to whether they hear /b/ or /p/, when there is no voicing left before the release burst. This means that voicelessness is not sufficient to identify the sound as [p]. The VOT concept contributes nothing to the explanation of this fact, but the IPA phonetician's classification in terms of lenis and fortis plosives does.

Moreover, IPA phonetics not only makes concrete statements about the realization of phonological word contrasts which measurements, such as VOT, cannot always capture, it can also provide insightful descriptions of communicative aspects of speech. In the early 1960s the French comedian Fernand Raynaud had his audiences in stitches with his explanation of Scottish haggis, which he called «La panse de brebis farcie», pronounced [la pʰãs də bɹɑbi fɑɹsi], in IPA transcription. He transmitted the ironic meaning of «spécialité de la cuisine anglaise» by adding certain sound properties of an English accent in French, especially the very long and strong aspiration. What he did to his vocal apparatus over time to produce these sound effects is irrelevant because the essential sound aspects are clearly represented in the universal segmental transcription, and it is them that triggered roars of laughter in French audiences. But these essential phonetic properties go beyond what a language-specific phonemic transcription could convey.

On the other hand, if a phonetic statement aims at providing an economical orthography for a language, a minimal phonemic transcription is called for, based on a monosystemic set of phonemes, for instance in pronouncing dictionaries, especially for languages whose codified orthographies are highly irregular in their letter-sound relationships, such as English. The subtitle of Pike's *Phonemics – A Technique for*

Reducing Languages to Writing – captures this goal. However, if the aim is, for example, the parametric text-to-speech synthesis of spoken text, the multilayered temporal dynamics in the generation of speech, following on from a grapheme-to-phoneme conversion, are of prime relevance.

To put the VOT concept into proper perspective, we must not forget that it was developed by speakers of English on the basis of the contrast in English plosives that was phonologically categorized as voiced/voiceless, because orthography and IPA-type broad phonetic transcription distinguished b/p, d/t, g/k [Lisker and Abramson, 1964]. At the same time, the Haskins researchers discovered, in their analysis of spectrograms, that phonologically voiced stops were often not phonetically voiced at all, and phonologically voiceless stops showed greater time lag of voice onset after stop release than phonologically voiced ones. And in perceptual tests they also found that English listeners partitioned a voice-onset continuum for the two phoneme categories of voiced and voiceless, with a sharp boundary at about 20 ms of voicing lag [Abramson and Lisker, 1970]. This meant that the phonological categories voiced/voiceless could be projected onto this VOT continuum. The VOT continuum was also extended to negative values preceding the release point, and became a scale onto which different languages project their phonological voiced/voiceless distinctions at different points, e.g. in French the contrast is between negative VOT and short lag, 0–20 ms. In Hindi, the scale is partitioned into three phonology-related sections, negative VOT, short and long positive VOT. But in Korean, the threefold occlusion-release contrast cannot be linked to this single duration scale, a lenis/fortis contrast in forming and releasing the plosive occlusion needs to be considered.

Thus, VOT provided a phonetic measure for a phonemic contrast that continued to be called voiced/voiceless, although a phonation contrast during the occlusion was no longer a defining property in languages like English. The persistent use of voiced/voiceless in two different senses, phonological and phonetic, leads to constant confusion and misinterpretation of data. A recent example of this is the work on tone and voicing in Kera [Pearce, 2009], a UCL PhD thesis [Pearce, 2007], which L. comments on quite extensively. Westerners developed a script for Kera which distinguishes b/p, d/t and g/k as categorized by Western ears. Kera has three lexical tones, of which the high tone is never combined with b, d, g, the low tone never with p, t, k. In the English phonological tradition, Pearce [2007] identified the two sets /b, d, g/ and /p, t, k/ with a voicing contrast, i.e. with voiced and voiceless plosives, respectively, although /b, d, g/ never have phonetic voicing during the occlusion. And following the Haskins analysis of plosives, she also projected the Western categorization of /b, d, g/ and /p, t, k/ on the VOT continuum, ranging from 0 to 50 ms, but found extreme variability in VOT across all stops and covariation of VOT with tone. Her conclusion, rightly, is that VOT is not distinctive in Kera and that there is only a tone distinction. VOT variability has nothing to do with a voicing contrast but is linked to the production of tone and heightens tone perception. However, L.'s argument that this example provides a 'challenge to any notion of universal categorization' (*Hb*, p. 360), showing the Eurocentric basis of the universal categorical taxonomy of the IPA (*Hb*, p. 359), goes in the wrong direction. Well-trained IPA phoneticians would not have applied the phonological voicing category, but would have classified what they heard as [p, pʰ], [t, tʰ], [k, kʰ] and would have stated the associations of these with different tones – never high after [ptk], never low after [pʰ, tʰ, kʰ] – and would then have proceeded to analyse the acoustic properties of the different plosive – tone combinations with reference to

general phonetic aspects of speech production. Nowhere would a universal voicing taxonomy have been invoked.

The Kera example proves to L. that the universal categorization assumption of systematic phonetics cannot be valid. But a modified form seems feasible 'in which phones are defined in language-specific quantitative terms as a mean value on some measurable phonetic scale (or, more generally, as a central value in some quantitatively definable phonetic space, such as the vowel space defined by the value of the first two formants). That is, we could give up the idea of *universal* categorization, but still maintain the segmental idealization and still maintain the idea that the output of the phonology is a string of systematic phones which are then passed on to physical phonetics for realization. Such realizations could be quite variable without upsetting the quantitative definition of the phone. … language-specific allophones can be defined quantitatively, each with its own portion of phonetic space, as long as the overlapping distributions are statistically distinct' (*Hb*, p. 361).

This argument is void since the general phonetic voiced/voiceless categorization should not have been applied to the data at all. Moreover, IPA classificatory categories are auditory in the first instance, cf. Trubetzkoy [1939], pp. 13f:

> …the auditory and tactile training, which a good 'ear phonetician' has to undergo, consists in getting used to listening to sentences and words, or to tactile checking while pronouncing them, without paying attention to their meaning, rather to observe only their sound and articulation (*die Schulung des Gehörs und des Tastsinns, der sich ein guter «Ohrenphonetiker» unterziehen muß, besteht eben darin, daß man sich gewöhnt, Sätze und Wörter abzuhören oder beim Aussprechen abzutasten, ohne auf ihren Sinn achtzugeben, vielmehr nur ihre lautliche, bzw. artikulatorische Seite wahrzunehmen…*).

And their relation to, e.g., acoustic measures is a matter of experimental perceptual analysis, not of two-dimensional 'cloud' graphs, and inferential statistics, of acoustic production data.

But then L. goes one step further and also argues against the segmental idealization assumption in a discussion of the acoustic realization of unstressed vowels in minimal pairs like *Rosa's* and *roses* in a study of American English by Flemming and Johnson [2007]. The two vowels are phonologically distinct as [ə] and [ɨ] – schwa and barred i in traditional notation of American English since Trager and Smith [1957]. A scatter plot of the first two formant frequencies of 5 minimal pairs, each pronounced twice by 9 female speakers, shows considerable overlap of the values for the two vowels, which L. interprets

> that [ɨ] is essentially a subset of [ə].[2] There is no obvious way to reconcile this kind of distributional fact with a traditional phone-based transcription. … the distribution of the vowel in *Rosa's* appears to occupy a continuous space on the plot, not two separate spaces corresponding to two different transcriptions. That is, the quantitative data justify the statement that there are two distinct unstressed *phonemes* /ɨ/ and /ə/ in American English, but not that American English phonetic realizations allow us to distinguish two *phones* [ɨ] and [ə] occupying reasonably distinct areas of phonetic space (*Hb*, p. 361).

There are not sufficient grounds for this verdict. The scatter plot conflates: (1) data from 9 female speakers, whose F1/F2 have not been normalized but may vary a good deal, (2) highish front and non-high back vowels in the first syllables of the 5 pairs (*leases/Lisa's*, *hinges/ninja's*, *ages/Asia's*; *roses/Rosa's*, *rushes/Russia's*), which are

---

[2] In the reproduced graph, L. wrongly links the scatter plot only to *Rosa's/roses*, although it represents 5 minimal pairs.

bound to have an influence on the opening of [ə], (3) different intervening consonants [z, s, ʃ, dʒ], which can again affect F1/F2 of [ə]. The results of the ANOVAs show these influences quite clearly: for both F1 and F2 the factors Subject and Pair as well as the interaction Subject × Vowel are significant. So, it would have to be ascertained by experiment whether in a general IPA auditory assessment, in relation to the auditory reference points of the Cardinal Vowels, these speaker and context variables lead to different phone classifications, or whether the observer disregards them. L. makes the mistake of considering raw acoustic data the representative phonetic variation of phonemes. He should realise that there are the articulatory, acoustic and perceptual levels of manifestation of speech acts, and different articulations may result in the same acoustic output, and different acoustic output may lead to the same percept in phonetic assessment, outside phonemics.

L. sums up his argument against *systematic phonetics* and the *phone* being concepts of a general phonetic theory as follows: '… twentieth century theories of phonology were universally built on the assumption that phones and phonetic transcriptions are a scientifically appropriate language-independent representation of speech'. 'The supposed reality of phones was crucial to the role played in traditional definitions of the phoneme' but ' … the identification of the phones on which we base our theoretical definition of the phoneme is specific to a given language' (*Hb*, p. 353). This is certainly true of a great many phonemic solutions by phonemicists without proper training in IPA-type phonetic analysis. As was pointed out above, in connection with the voicing contrast and its projection on the VOT continuum, the language-specific orientation arises when phonological oppositions are given language-based abstract phonetic labels and are then filled post hoc with instrumental, usually acoustic, data, instead of working from systematic phonetic classification to phonemic categorization and phonetic measurement.

Distinctive Features

L. maintains that 'the idea of a universal scheme of classification for phones gives rise to what is perhaps the central theoretical construct of mid-20th century phonology, namely the feature'. His views on Trubetzkoy's phonologically relevant sound properties, determined by the systemic oppositions a phoneme enters into, as well as the bundling of these properties in the phonological essence of a phoneme, have already been discussed in 'Trubetzkoy's Two Sound Levels'. This concept of phonological sound oppositions allows the setting up of proportions, such as /pb/ : /td/ : /kg/ or /bm/ : /dn/ : /gŋ/ in German, in which pairs of phonemes are characterized by the same distinctive mark. Jakobson, Fant and Halle [1952] (*JFH*) developed this into their *distinctive feature* theory, which L. characterizes as follows:

> The most conspicuous taxonomic innovations were that the features were exclusively binary and that they were defined in purely acoustic terms. However, the *JFH* feature system reaffirms the two premises of systematic phonetics …: it presupposes the segment, and it explicitly presents the taxonomic framework as universally valid. … the *JFH* version of the segmental idealization does represent a refinement of the IPA version, … it treats the segment … as an idealized instantaneous slice through the signal at a specific point in time … The features that characterize the segment are therefore based on the acoustic properties of the signal at the point in time when the idealized instantaneous slice is taken (*Hb*, p. 356).

To say that the *JFH* distinctive feature system is 'purely acoustic' misrepresents it. *JFH* were concerned with the transmission of phonological information, contained in

phoneme features, through a sequence of phonetic manifestations from articulation to articulatory acoustics to signal acoustics to aural processing to perception. The closer the investigation is to perception by the receiver, the more accurately can the information conveyed by its sound shape be gaged. 'This determines the operational hierarchy of levels of decreasing pertinence: perceptual, aural, acoustical and articulatory… The systematic exploration of the first two of these levels belongs to the future …' (*JFH*, p. 12). This comprehensive concept of distinctive features and of their manifestations goes back to Roman Jakobson, for whom the auditory aspect was most important [cf. Fischer-Jørgensen, 1995, p. 152f.].

The *JFH* distinctive feature system is thus integrated into a framework of speech communication, which distinguishes it from IPA classification and transcription as well as from traditional phonemics and later generative phonology, and as a communicative system, it builds on universals in speech production and perception of *homo loquens*:

> …the specification of the phonemic oppositions may be made in respect to any stage of the speech event from articulation to perception and decoding, on the sole condition that the variables of any antecedent stage be selected and correlated in terms of the subsequent stages, given the evident fact that we speak to be heard in order to be understood (*JFH*, p. 13).

And since the systematic perceptual categorization of the features was a task for the future, the acoustic level was closest to the primary perceptual level and therefore provided the domain for defining the *JFH* feature system, 'articulatory … carrying no direct information to the receiver' (*JFH*, p. 12; Jakobson did not believe in the motor theory of speech perception).

But even at the acoustic level we have to differentiate the acoustics of source and transfer function in the vocal tract from the acoustics of signals radiated from the lips. *JFH* are concerned with the former, relying on three-dimensional frequency-energy-time spectrographic acoustic signal analysis by the hardware sonagraph, which gives energy distributions in frequency regions that can be related to properties of the source and cavity resonances in an acoustic theory of speech production [Fant, 1960]. The *JFH* distinctive feature framework certainly has its weaknesses at various points, as regards their definitions, their universal fixation to 12, their strict binary nature, but to specify their manifestations on all the levels from articulation to perception within a framework of communication between speaker and listener is a task at which phoneticians, particularly in perception research, have been working over the past half century, and which therefore cannot be brushed aside as taxonomic classification of phones.

In Chomsky and Halle [1968], the application of distinctive features went in a different direction. Since the level of taxonomic phonemics was abandoned, and morphophonological relations were incorporated in *classificatory* matrices of lexical entries, the features in these could no longer be physically defined. They continued to be strictly binary, but assumed abstract specifications in spite of phonetic, now articulatory, labels. The function of the phonological component then consists in deriving *phonetic* representations (matrices) by rules operating on classificatory features of lexical items in a surface structure assigned by the syntactic component. It is only this phonetic-matrix output that receives a physical interpretation of its sequential segments, with the same feature labels, which may now be graded. So, for instance, word-final obstruents in German (*bunt, Bund*) are specified as [±voice] in the underlying classificatory matrix because of their paradigmatic alternations (*bunte, Bünde*), although they are physically [–voice] in their phonetic feature.

This use of the same label *voice* for the classificatory and the phonetic function of the distinctive feature leads to constant confusion and has triggered a lot of post hoc experimentation, where the phonological desk solution is taken to the laboratory for phonetic validation by measurement, following the line of thought 'if there is no underlying phonological neutralization, there may be no phonetic neutralization either'. None of these investigations hold water [Kohler, 2012], for methodological reasons of data collection, for reasons of values being below the just noticeable difference, and not explaining the facts of, e.g., German-accented English. Generative phonology replaced the communicative perspective of information transmission from speaker to hearer in *JFH* with the cognitive representation of language in the ideal speaker/hearer. Such phonological solutions should *not* be taken into the laboratory to be filled with (acoustic) parametric substance, but they *should* be tested as to their validity by experiment in phonetics.

Formal Model of Phonology

The Phonology-Phonetics Interface

Trubetzkoy divided the field of the sound of language and speech into phonology and phonetics, the latter being, first and foremost, the time courses of signal parameters of early instrumental phonetics. The question then arose, in the wake of generative phonology, as to how symbolic discrete representations are transformed into parametric continuous ones in phonetics. This created the notion of a phonology-phonetics interface, bringing the separated fields together again, and led to the discussion of where exactly this interface is to be placed in the speech chain. This issue is also important in L.'s discussion of the place of systematic phonetics:

> Systematic phonetics is often seen … as a level of representation at the interface between the abstract and the physical. … any detailed scientific description of physical processes must eventually be expressed in quantitative parametric terms. … the level of description beyond which continuous parametric models are required is the phonemic representation, and … all other details of utterance phonetics cannot usefully be described in terms of symbolic categories (*Hb*, pp. 361–363).
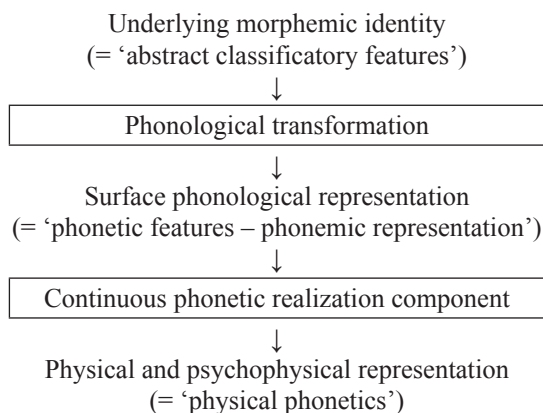
Symbolic idealization at the systematic phonetic level of description is thus ruled out as an adequate scientific account of the sound of language. This conclusion is supported, so the argument goes, by the facts of assimilations across word boundaries. In generative phonology they are treated as phonetically complete in the generated phonetic matrix although they may be developing over time. But L. maintains that these processes should not be incorporated in the phonetic matrix but should be part of the parametric continuous phonetic component, which in turn means that the interface occurs at the phonemic level, and the systematic phonetic level is no longer necessary. Or as Ladd [2006, p. 12] puts it: 'Chomsky argued that various problems arise from positing two intermediate symbolic representations [phonemic and systematic phonetic], and proposed to do away with the taxonomic phonemic level … however, we can see that they eliminated the wrong one.' So, we are back to the traditional phonemic representation with partially specified phonetic segments, preceded by an abstract phonological representation and followed by continuous realization in a physical phonetic representation.

L. connects the physical phonetic representation to the more recently established field of enquiry of *laboratory phonology*. '… the phrase is precise and meaningful: laboratory phonology examines the sounds of *parole* not in order to learn more about

the processes of speech production and perception, but to evaluate the implicit predictions that phonological representations make about phonetic behavior' (*Hb*, p. 358). Laboratory phonology takes systematic phonological desk solutions into the laboratory for measurement and substantiation, leading to (instrumental and experimental) *phonetics in phonology*.

Levels of Representation and 'Phonetics in Phonology'
The result is a phonological grammar in a formal cognitive model of a language, which may comprise the following levels of representation [cf. also Ladd, 2006]:

Underlying morphemic identity
(= 'abstract classificatory features')
↓

Phonological transformation

↓
Surface phonological representation
(= 'phonetic features – phonemic representation')
↓

Continuous phonetic realization component

↓
Physical and psychophysical representation
(= 'physical phonetics')

The details of such a model are, however, far from clear. For a start, L. uses the term *representation* 44 times in a variety of senses:

–   The most frequently used phrasing 'systematic phonetic representation'/'detailed (but still symbolic and segmental) phonetic representation' occurs in the extensive discussion of IPA symbolization. It refers to the IPA-type representation of speech, which L. rejects as invalid in his cognitive model of phonology. The phrasings 'the IPA idea that there is a universally valid segmental representation of utterances in any language' and 'the assumption that phones and phonetic transcriptions are a language-independent representation of speech' belong to the same discussion context.
–   In the phrasing 'a faithful representation of what a speaker really produces', it is again speech events that are being represented, but their representation is not faithfully specified. Is it the articulatory dynamics, the acoustic consequences of articulation in the acoustic signal, or the physiological processes leading to articulation? Moreover, if IPA phoneticians represent English 'bat'/'pat' as [b̥ætʰ]/[pʰætʰ], French 'batte'/'patte' as [bat̪]/[pat̪], they quite faithfully represent the essential aspects of the speakers' speech events, especially the occlusion-release-opening sequence at the beginning of each word. So, what characterizes a faithful representation?
–   'Symbolic segmental representation', 'symbolic or discrete representation' refer to the representation of elements of *langue* as against the physics of *parole*, but what is actually represented in *langue*, and what is its ontological status?
–   The phrasings 'phonological representations', 'abstract "systematic phonemic" representations', 'phonemic representation', 'interface representation', 'underlying

representations' refer to various types of representation of elements in *langue* in relation to various ways of cognitive modelling of phonology. The status of what they represent is defined by the theoretical framework within which the modelling is done, there is thus circularity between the represented and the representation.

The levels of representation in L.'s model of phonology are built on the strict separation of *langue* versus *parole* and of discrete grammar-related phonology versus continuous physical phonetics, and on the interfacing of the dichotomous domains. L.'s model stands in the tradition of generative phonology and the cognitive orientation of its theoretical constructs. The way L. differs concerns a reorganization of the levels of representation and the integration of the continuous phonetic level into the phonology for phonetic evaluation of phonological representations. The levels of representation are projected onto the field of cognition and thus invested with reality for the speaker/hearer. The progression from abstract phonology to concrete phonetics, rather than the reverse, is thus theoretically pre-established. Phonologists arrive at their phonological systematizations of symbolic phonetic language data by deduction from their theoretical premises without independent experimental testing, and then take them into the laboratory for phonetic substantiation.

Answering the Three Questions Asked at the Outset

(1) Does the historical survey provide a faithful account of phonetic concepts in relation to phonology? I think it has become clear that the historical survey of IPA phonetics, Prague phonology and *JFK* distinctive feature theory is not a faithful account.

(2) What type of formal model does L. propose for phonology? It derives representations at a sequence of levels from abstract symbolic morphophonology to continuous physical time courses, and it is a model in which the various representations are assumed to have cognitive status. It thus aims at providing a cognitively explanatory description of how language and speech are structured ('language as a cognitive system', cf. Pierrehumbert et al., [2000]). Speech scientists' ultimate goal must certainly be to elucidate the cognitive anchoring of speakers' and hearers' behaviour, but it is by no means straightforward that phonological representations within a model like L.'s can be projected onto the cognitive level and invested with reality for speakers and hearers. The German speaker/hearer will no doubt have internalized a relationship between *Kind* [t] and *Kinder* [d], but to say that the underlying representation for [t] in *Kind* as [+voice] is more than a device of descriptive systematization, and captures a cognitive structure, is an unfounded hypostatization, as long as there is no experimental evidence from outside the model. The same kind of argument applies to all other cases of positional phonemic neutralization in various languages, and to [ʃ] of English *impression* and *separation* having underlying representations as /s/ and /t/, respectively, because of *impress* and *separate*.

Like most work in theoretical phonology in the last 50 years, L.'s model excludes the important communicative function of speech and language in speaker-hearer interaction in situational settings. It is thus not a process model of speech communication. However, the central questions speech scientists should be asking are 'How do humans *communicate* with speech in the languages of the world, and are there

Kohler

common – typological and universal – *communicative* patterns across languages?' L.'s model is not oriented towards answering these core questions.

(3) Is 'Phonetics in Phonology' a convincing solution to the phonology-phonetics dichotomy? Is it appropriate to absorb *phonetics* into *phonology*, considering that an adequate evaluation of sound patterns of *homo loquens* needs to take the primacy of relationships between speaker, listener and communicative situation into account, over and above, and independently from, the sum of the formal relations?

When in the early 20th century, the Prague School developed phonology as a linguistic study of sound there were two fields of phonetics: (i) IPA descriptions of the sounds of languages, combining, within a framework of linguistic meaning, two aspects of sounds – those used to differentiate words and those that add phonetic detail, and (ii) the newly established instrumental phonetics, making records of speech events outside linguistic meaning. Two different methodologies defined these fields – the humanistic approach of sensory observations and their systematic descriptions, and the science approach of instrumental measurement. The Prague School linked these to de Saussure's dichotomy of *langue* and *parole*, and at the same time developed (i) into a study of sound systems, based on distinctive difference and contextual variation, and called it phonology. Trubetzkoy renamed the two fields as 'sound in language systems' (*Sprachgebildelautlehre*) and 'sound in speech acts' (*Sprechaktlautlehre*). The two fields were conceived of as being so fundamentally different that they represented two different domains of study. This also entailed that a general phonetic classification, for example in field work before the sound system can be worked out, became part of *Sprechaktlautlehre*, although it was segmental, not a physical parameter over time.

The two fields of study were procedural domains for making different sets of statements about sound in language and speech. But Trubetzkoy was also very clear that *Sprechaktlautlehre*, as far as general segmental classification was concerned, was a procedural prerequisite to *Sprachgebildelautlehre*. As the Prague School developed phonology in the humanities as a counterpart to phonetics in natural science, it was not concerned with relating the two. Their relationship and the integration of phonetic measurement into linguistic analysis became the central issue raised by Eberhard Zwirner in his "Phonometrics" (*Phonometrie*) [Zwirner and Zwirner, 1936] where various levels of statements from segmental phonetic classification and phonemic categorization to statistically evaluated segments of speech signal traces are *related* to each other (their term is *Zuordnung*). They do not form a linear cognitive sequence but are parallel exponents of phonetic events in related statements.

The question of how sound systems materialize in speech in a model of phonology or of speech production did not arise. The first time the issue of the transformation of units across a sequence of levels entered the discussion was in the distinctive feature theory of Jakobson et al. [1952]. But there the levels were different phonetic manifestations in speaker-hearer communication, and since the focus was just on the acoustic level, modelling of a linear sequence of representations did not become an issue. This changed with generative phonology when units were transformed from an abstract underlying phonological representation to a symbolic phonetic surface representation in a phonological component of a grammar of the language, and these levels received cognitive status. The physical phonetic signal representation stayed outside phonology, thus maintaining the phonology-phonetics dichotomy. L. has removed the symbolic phonetic surface representation of generative phonology, has reinstated the classical phonemic level, and has also incorporated the continuous phonetic level

into phonology. In principle, nothing has changed, the levels of representation, and especially the phonology-phonetics interface, have been shifted, but the problems connected with the cognitive status of the levels of representation and the discrete-to-continuous transformation persist. The latter problem is now to be solved by filling pre-established phonological representations with phonetic measurement data in laboratory phonology.

An alternative framework for the analysis of sound in speech and language gives up modelling phonology as a linear sequence of cognitive levels of representation, and returns to Trubetzkoy's idea of making statements in different methodological environments, and then takes up Zwirner's phonometric approach relating these different sets of statements to gain insight into the working of speech communication. This in turn means abandoning the phonology-phonetics division in a *communicative phonetic science*, based on the analysis of speech from articulation via acoustics to perception in the languages of the world. It includes questions about phonological systems and structures, not only of the type raised by the Prague School, but by Firthian Prosodic Analysis as well [Ogden and Local, 1994], thus giving us *phonology in phonetics* instead of *phonetics in phonology*. And in such a phonetics paradigm, phonological solutions come out of the laboratory through experimental phonetics in a communicative framework, rather than going into the laboratory for phonetic substantiation, as in laboratory phonology [Kohler, 2007]. This paradigm will be able to give insightful answers to the core questions raised under (2) above, and contribute to the modelling of speech communication.

Acknowledgement

References

Abercrombie, D.: English phonetic texts (Faber & Faber, London 1964).

Abramson, A.S.; Lisker, L.: Discriminability along the voicing continuum: cross language tests. Proc. 6th Int. Congr. Phonet. Sci. 1967, pp. 569–573 (Academia Publishing House of the Czechoslovak Academy of Science, Prague 1970).

Baltaxe, C.: Principles of phonology (translation of Trubetzkoy's Grundzüge, 1939; University of California Press, Berkeley 1969).

Bell, A.M.: Visible speech, the science of universal alphabetics; or self-interpreting physiological letters, for the writing of all languages in one alphabet (Simpkin, Marshall & Co., London; Trübner, London 1867).

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

Bloch, B.: Studies in colloquial Japanese IV: phonemics. Language *26*: 86–125 (1950).

Brücke, E.W.: Grundzüge der Physiologie und Systematik der Sprachlaute für Linguisten und Taubstummenlehrer (Gerold's Sohn, Wien 1856).

Bühler, K.: Phonetik und Phonologie. Travaux du cercle linguistique de Prague *4*: 22–53 (1931).

Cantineau, J.: Principes de phonologie (translation of Trubetzkoy's Grundzüge, 1939; Klincksieck, Paris 1949).

Chomsky, N.; Halle, M.: The sound pattern of English (Harper & Row, New York 1968).

Fant, G.: Acoustic theory of speech production (Mouton, 's-Gravenhage 1960).

Flemming, E.; Johnson, S.: *Rosa's roses*: reduced vowels in American English. J. Int. Phonet. Ass. *37*: 83–96 (2007).

Fischer-Jørgensen, E.: Trends in phonological theory until 1975. A historical introduction. Travaux du cercle linguistique de Copenhague, vol. XXVII (Reitzel, Copenhagen 1995).

Goldsmith, J. (ed.): Handbook of phonological theory (Blackwell, Oxford 1995).

Hawkins, S.; Smith, R.: Polysp: a polysystemic, phonetically-rich approach to speech understanding. It. J Ling. *13*: 99–188 (2001).

Hocket, C.F.: A system of descriptive phonology. Language *18*: 3–21 (1942).

IPA: The principles of the International Phonetic Association (Department of Phonetics, University College London, London 1949).

IPA: Handbook of the International Phonetic Association. A guide to the use of the international phonetic alphabet (Cambridge University Press, Cambridge 1999).

Jakobson, R.; Fant, C.G.M.; Halle, M.: Preliminaries to speech analysis. The distinctive features and their correlates (MIT Press, Cambridge 1952).

Joos, M.: Acoustic phonetics. Language *24*: suppl. 2 (1948).

Kohler, K.J.: Contrastive phonology and the acquisition of phonetic skills. Phonetica *38*: 213–226 (1981).

Kohler, K.J.: Beyond laboratory phonology: the phonetics of speech communication; in Solé, Beddor, Ohala, Experimental approaches to phonology, pp. 41–53 (Oxford University Press, Oxford 2007).

Kohler, K.J.: NEUTRALIZATION.?! The never-ending laboratory phonology issue of word-final obstruent voicing; in Gibbon, Hirst, Campbell, Rhythm, melody and harmony in speech. Studies in honour of Wiktor Jassem, pp. 171–180 (Polish Phonetic Association, Poznań 2012).

Ladd, D.R.: 'Distinctive phones' in surface representation; in Goldstein, Whalen, Best, Lab. Phonol., vol. 8, pp. 3–26 (de Gruyter, Berlin 2006).

Lisker, L.; Abramson, A.S.: A cross-language study of voicing in initial stops: acoustical measurements. Word *20*: 384–422 (1964).

Mitchell, T.F.: Colloquial Arabic: the living language of Egypt (English Universities Press, London 1962).

Ogden, R.; Local, J.K.: Disentangling autosegments from prosodies: a note on the misrepresentation of a research tradition in phonology. J. Ling. *30*: 477–498 (1994).

Pearce, M.: The interaction of tone with voicing and foot structure: evidence from Kera phonetics and phonology; diss. University of London, London (2007).

Pearce, M.: Kera tone and voicing interaction. Lingua *119*: 846–864 (2009).

Pierrehumbert, J.; Beckman, M.E.; Ladd, D.R.: Conceptual foundations of phonology as a laboratory science; in Burton-Roberts, Carr, Docherty, Phonological knowledge. Conceptual and empirical Issues, pp. 273–303 (Oxford University Press, Oxford 2000).

Pike, K.L.: Phonetics: a critical analysis of phonetic theory and a technique for the practical description of sounds (University of Michigan Press, Ann Arbor 1943).

Pike, K.L.: Phonemics: a technique for reducing languages to writing (University of Michigan Press, Ann Arbor 1947).

Sievers, E.: Grundzüge der Lautphysiologie (Breitkopf & Härtel, Leipzig 1876).

Sweet, H.: A handbook of phonetics (Clarendon Press, Oxford 1877).

Trager, G.L.; Smith, H.L.: An outline of English structure. Stud. Ling.: occas. Papers 3 (American Council of Learned Societies, Washington 1957).

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

Whitney, W.D.: A Sanskrit grammar (Breitkopf & Härtel, Leipzig/Kegan Paul, Trench, Trübner, London 1896).

Zwirner, E.; Zwirner, K.: Grundfragen der Phonometrie (1939); 2. Aufl. (Karger, Basel 1966).