The Whole Theory of Sound Structure

Janet Pierrehumbert

Department of Linguistics, Northwestern University, Evanston, Ill., USA

Introduction

The theory of phonology describes implicit knowledge that people have about speech sound structure as it functions in language. The concepts of phonology, as attributed to speakers and listeners, make it possible to relate speech sounds and articulations, on the one hand, to contrasts in meaning, on the other. It is hardly possible that phonological concepts are unrelated to the physical reality of speech as studied in phonetics. Indeed, following Quine [1964], I would suggest that even the most abstract ideas of physical science and higher mathematics arise from our efforts to interpret and predict physical experience. The language learner may be viewed as a baby scientist, who in acquiring the phonology of his language constructs a theory of speech as it functions to distinguish meanings; the concepts in this theory are therefore 'about' speech.

The phonological concepts that linguistic theory attributes to speakers, such as the syllable or the feature [coronal], are among the more concrete of the concepts that we take to populate the human mind. They are certainly far more concrete than commonplace ideas like 'good' or 'reincarnation'; broadly speaking they are on a par with concepts like 'red', 'heavy', or 'kitchen'. Accordingly, I believe that their relationship to physical reality is particularly intimate, even if it is still in many ways perplexing to scientists. It is a pleasure to observe the degree of consensus on this point which is found in the papers for this volume, particularly in view of the past history reviewed by Diehl [this vol.].

Phonetic Implementation

One result of this emerging consensus has been a productive line of research concerning the division of labor between rules of phonology and principles of phonetic implementation. In the framework of generative phonology, rules of phonology specify what combinations of phonological categories are well formed. Above all, they characterize 'possible words', but they also characterize e.g. 'possible intonation patterns' insofar as these are categorical and function conventionally to convey meaning.

The phonological rules accomplish this task via their aggregate activities within a single type of representation, which is a qualitative one and which is, speaking technically, in the purview of formal language theory. Rules of phonetic implementation, by contrast to phonological rules, relate the categorical representations of phonology to representations of a completely different character – quantitative descriptions of speech as a physical phenomenon. By doing so, phonetic implementation rules specify what phonological representations mean, a point developed at more length in Pierrehumbert [1990].

The division of labor between phonological rules and principles of phonetic implementation is often unclear (just as elsewhere in linguistic theory the division of labor between syntax on the one hand and semantics and pragmatics on the other is often unclear). Aspects of sound structure which are regular but not distinctive may appear to be amenable to analysis in either way. A very detailed investigation is often required to determine which type of analysis is better.

In my understanding, the paper by Kohler [this vol.] argues that a class of facts about allophony in German is best handled through phonetic implementation rules rather than through the qualitative rules of phonology. The argument is based on the finding that the reduction processes are characterized by a continuum which phonological rules do not capture. This position is in broad agreement with the outcomes of a number of other studies dealing with the competition between principles of phonetic implementation and low-level phonological rules. Such studies have in general led to the conclusion that phonetic implementation rules are responsible for many regular but nondistinctive aspects of sound structure which might on slight examination appear to be categorical. These rules are powerful and because they are language-particular, we must believe that they are learned and mentally represented. As a result of such studies, the surface phonological description (the most superficial description of sound structure which is still qualitative rather than quantitative) is found to be more abstract than previously supposed. In other words, the categorical component of the theory of sound structure has less work to do than once thought. We may note that this development has been paralleled by a similar development on the other side of linguistic theory, where semantic and pragmatic principles are taking over more and more of the work previously assigned to the sentence syntax.

Let us consider some additional examples. Pierrehumbert [1980] showed that English intonation can be described using two tones (L and H) instead of the four previously proposed, provided that a nontrivial rule of phonetic implementation is posited. This rule, downstep, applies iteratively left to right, with the outcome at any particular point depending both on the tonal categories and on the immediately prior quantitative result. This characterization of the rule was supported in detail by the experimental results and model developed in Liberman and Pierrehumbert [1984]. In addition to simplifying the phonology and providing predictions about the phonetic outcome, the proposal handles phrases with more than four steps, which were problematic for previous theories which had four tones but trivial principles of phonetic implementation. One consequence of this work is that

the output of the phonology as conceived in Chomsky and Halle [1968] is not a well-defined level of representation; mapping the tones into n-ary tonal features would neither accurately represent the quantitative form of downstepped sequences, nor maintain the information needed to compute this form. Since languages are also known to differ both in the structural condition for downstep and in its extent, the work also implies that phonetic implementation rules are language-specific.

Selkirk [1984] proposed that phrase-final lengthening be treated phonologically, by adding silent beats at phrase boundaries. These silent beats are instantiated either as silence or as prolongation of preceding material. Results of the timing study in Silverman and Pierrehumbert [1990] show that this proposal is inadequate, since it fails to describe both the location of lengthening and the way in which different factors which induce lengthening interact. They suggest that the mechanism which controls timing instead refers directly to the phrasing and other prosodic determinants of duration.

Fourakis and Port [1986] found that the so-called epenthetic /t/ in words like *tense* is not identical to the underlying /t/ in corresponding words such as *tents*. Furthermore, the existence of a silent gap between the nasal murmur and the fricative in *tense* was found to be dialect-dependent. Their results indicate that there is no phonological rule of /t/ epenthesis. Instead, a dialect-specific rule of phonetic implementation controls the timing of velum motion. A pseudo stop gap results when the velum is closed before the tongue blade releases into the fricative. Browman and Goldstein [1990] also discuss examples in which prin-

ciples of timing plausibly explain observed allophony.

There have also been a number of proposals by phonologists which critically depend on the availability of powerful phonetic implementation rules. I mention two such proposals in order to give an idea of the change that has occurred within generative phonology since Chomsky and Halle [1968].

Hayes [1984] discusses some apparent counterexamples to the rules governing voicing assimilation in Russian. Some sonorants which are in a position to be devoiced are actually voiced. Hayes [1984] suggests that the phonetic implementation rules are responsible because a laryngeal configuration which may effectively suppress vocal fold oscillation in an obstruent may permit it in an adjacent sonorant, which has a more open vocal tract configuration. That is, he proposes that the sonorants in question are phonologically [−voice], but show up with a periodic source anyway. Kiparsky [1985] adopts this suggestion. A detailed phonetic study (for example, observations with a stereo fibroscope) would be needed to evaluate this suggestion.

Lombardi [1990] surveys phonological rules referring to affricates, in order to assess whether affricates are phonologically contour segments (first [−continuant] then [+continuant]). If they are contour segments, affricates should group with stops when appearing at the right edge of the structural description for a rule, and with fricatives when appearing at the left edge. The observed asymmetry is not found. Affricates may form a natural class either with stops or with fricatives, regardless of where they appear in the structural description. This result leads Lombardi [1990] to propose that affricates are phonologically

[+stop +continuant], with the temporal ordering of the frication relative to the stop gap established after the phonological rules. Rules governing (nondistinctive) aspiration or the release of stops, which can be sensitive to the temporal ordering of the closure and fricative regions, are taken to be obvious examples of phonetic implementation rules.

Since phonetic implementation rules are both powerful and language-particular, it is (strictly speaking) unclear that the $\tilde{V}$NC-[+voice] sequences analyzed by Ohala and Ohala [this vol.] have a nasal segment. The reason provided for suggesting that the nasal is phonologized here (but not in $\tilde{V}\#$C[+voice]) is that the nasal murmur is shorter before a word boundary than word-internally. However, nothing we know about phonetic implementation actually prevents the formulation of a rule that exploits the structural position to generate the observed murmur durations directly from a $\tilde{V}(\#)$C[+voice] representation. Many researchers might assume, with Ohala and Ohala [this vol.], that such a rule is unlikely – a longer murmur would be expected before the word boundary than internally. However, this assumption is not firmly supported in the absence of a general quantitive model of duration for Hindi. My general point is that, given the great encroachments of phonetic implementation rules on phonological turf, we need to think carefully about what evidence we view as sufficient to show that an observed pattern is phonological rather than phonetic.

I would also like to comment on Blumstein's [this vol.] discussion of German final devoicing in light of the perspective just laid out. Blumstein's conclusions appear quite startling. In the understanding of most researchers, a phonological rule which converted [+voice] segments to [–voice] ones would result in complete neutralization. The output for words with underlying voiced and voiceless obstruents would be identical, both for subsequent phonological rules if any, and for the phonetic implementation rules. The understanding that phonological representations and rules are strictly categorical is in fact crucial to work like Lombardi's [1990], which builds on the finding that certain information (in this case, the timing pattern within affricates) is *unavailable* within the phonology. If phonological representations were phonetically interpreted at every level of abstraction, then we would expect that any regular phonetic property of any segment could function phonologically.

Therefore, any residue of phonetic contrast between underlying voiced and voiceless obstruents provides evidence that a phonological rule has not applied. The fact that the contrast between voiced and voiceless obstruents is only poorly marked in word-final position would be attributed to the phonetic implementation of voicing as a function of prosodic position. If the acoustic contrast is slight, in comparison to the statistical variability, then it is to be expected that the contrast will be difficult or impossible to acquire. In fact it appears that word-final voicing in German is not learnable in the absence of a supporting morphophonemic alternation, and so word-final voiced obstruents have been lost historically on stems which cannot carry suffixes.

Blumstein [this vol.] does not agree with this summary of the situation, arguing that phonetic evidence of partial neutralization supports the existence of a phonological rule. Her position entails a kind of inter-

penetration of categorical and phonetic characterization which would surely be viewed as radical by most generative phonologists. However, the discussion in Diehl's essay [this vol.] suggests that he would be sympathetic to this idea. Fowler [1990] goes even further by strongly rejecting a division between categorical and quantitative characterization of language. Thus these three authors, whose perspectives are different in so many regards, all agree in raising the possibility that phonology is not as grammatical and categorical as the school of generative phonology would claim.

The possibility that the entities of phonology are not really categorical, but are rather quantitative, phonetically interpreted, and statistically variable, merits serious attention. However, it requires a kind of support which has not been provided. Proponents of this idea could advance their case by examining in detail rule classes which phonologists would view as demonstrably phonological – for example, rules which can be shown to be cyclic. They might attempt to demonstrate that such rules are sensitive to phonetic information which phonological theory fails to represent – in short, that phonologists are incorrect about what information is unavailable to rules of word information. In general, proponents of this view would need to provide positive and fully formulated proposals about what sequences of elements are well formed, both in language in general and in particular languages.

Some Issues in Phonology

Present phonological theory sets great store by the phonetic content of phonological representations. In some cases this has produced naive and reckless speculation which any researcher with laboratory experience would be unlikely to give credence to. In other cases, issues have been raised which are of deep importance to linguistic phonetics. The papers in the present collection take up just a few of the potential points of contact between phonetics and phonology. That is, these papers deal mainly with the phonetic theory of phoneme inventories or distinctive features (as in contributions by Beddor, Kingston, and Maddieson, and in Blumstein's discussion of stridency), and with the phonetic explanation or reanalysis of phonological rules (contributions by Kohler and Ohala and Ohala). Though these papers are all extremely interesting, I believe that a broader range of issues could be profitably addressed. I would like to sketch some of these issues here.

Feature Geometry

The theory of feature geometry, reviewed in this journal by McCarthy [1988], deals with the pattern of dependencies among distinctive features as observed in phonology. According to this theory, the distinctive features are organized into a tree structure on the basis of major articulators used to produce them. For example, the root node dominates a laryngeal node and a place node; features such as [spread glottis] or [voice] are subordinate to the laryngeal node, while features such as [labial] or [coronal] are subordinate [with further intervening structure] to the place node. One motivation for this hierarchical structure is the way features group in assimilations. For example, total assimilation of place is readily described by autosegmental spreading of

a place node with all its subordinate features. In the formalism of Chomsky and Halle [1968], total assimilation of place must be described using a complex rule which is formally no simpler than many unnatural rules would be. A second motivation for the theory of feature geometry is the fact that some features are only relevant for some classes of sounds. For example, [lateral] is only relevant for sounds in which an articulation of the tongue is defined, and so would only appear under a node which specifies activation of the tongue; if there is no such activation [as in /h/], this entire section of the tree would be missing.

The idea that some features are left unspecified is carried very far in much current work. For example, Lombardi [1991] argues that the only segments which can carry a distinctive feature for voicing are the voiced obstruents. All other segments are unspecified for voicing – a general process of interpretation yields the result that sonorants are voiced and all obstruents not marked for the feature [+voice] have no voicing. In fact, according to Lombardi [1991], in positions where laryngeal oppositions are neutralized (for example, if a language has plain, glottalized, and breathy stops, but permits only plain stops in syllable-final position), the neutral variant will have no laryngeal node at all.

Such an analysis actually undermines the interpretation of nodes in the feature tree as active articulators. For Lombardi [1991], it cannot be the case that the laryngeal node (with its dependents) phonologically specifies the physical state of the larynx. The neutral stop lacks a laryngeal node, but still is produced with a determinant laryngeal configuration which differs from the configuration for vowels.

Therefore, the physical state of the larynx is attributed to a broader evaluation of the tree, in particular the feature [sonorant].

Coming from a completely different direction, the papers by Kingston [this vol.] and Diehl [this vol.], also call into question the idea that features are organized into groups according to the major active articulator. These papers even lead one to wonder whether the pattern of co-occurrence and dependence among features can reasonably be reconstructed as a tree structure. Experimental work leading to positive proposals about the formal character of these dependencies could make an important contribution to the theory of sound structure.

Lexical Phonology

Lexical phonology [reviewed in Kaisse and Shaw, 1985 and Kiparsky, 1985] is a theory of the form and interaction of phonological rules. It emphasizes processes of word formation, drawing major implications from the claimed interleaving of morphological and phonological operations. It has built up a picture in which word level phonology is effectively constrained by the inability to refer to or create nondistinctive phonological categories. For example, if a language did not use place of articulation for nasals to distinguish words, then a lexical rule could not create a distinctively velar nasal. This is known as the principle of Structure Preservation.

Postlexical rules are not constrained by Structure Preservation. The English rule of aspiration for voiceless stops, for example, is described by most authors before Lombardi as postlexical rule introducing a nondistinctive feature (namely aspiration). The

rule could not be lexical because aspiration is not distinctive in English. Rules which apply across a word boundary are also clearly postlexical; the structural description for such rules could not be met in the lexicon, where words have not yet been assembled into phrases.

In the theory of Lexical Phonology, the lexical rules are highly constrained and the postlexical rules are freed from the lexical constraints. However, it is probably the case that postlexical rules conform to principles which do not necessarily govern lexical rules. I make this suggestion particularly because the present theory of lexical rules has (unnoticed by its inventors) drifted in the direction of a formal theory of analogy amongst words. This drift has occurred both within Lexical Phonology proper, and in other work on word formation within generative phonology, e.g. Halle and Vergnaud [1987]. Postlexical rules, in contrast, are highly productive, applying on the fly to arbitrary sequences of words, and generating output which either is phonetic or can serve as input to the phonetics.

There is at present no general theory of postlexical rules; only effective integration of work in phonology and phonetics can produce one. Here are some of the questions which such work might address: Do feature-changing postlexical rules exist, or are all candidates for such rules actually phonetic implementation rules, as suggested in Liberman and Pierrehumbert [1984]? What are the consequences of the fact that speech is produced and understood in time? Do any formal properties of postlexical rules follow from their high productivity? In general, what do postlexical representations look like and how are these representations manipulated?

Prosody, Suprasegmentals,
and Phonological Structure

Work on prosody and suprasegmentals has played an extremely central role in the development of phonological theory in the last 20 years. The initial development of metrical and autosegmental phonology was driven by the shortcomings in Chomsky and Halle's [1968] treatment of grouping, prominence, tone and intonation. The technical developments in these areas were then found to provide great improvements in the treatment of aspects of segmental phonology. They play a critical role in our present understanding of vowel harmony, assimilation, gemination and degemination, reduplication and phonotactics [Clements, 1981; Marantz, 1982; Hayes, 1986; McCarthy, 1986; Ito, 1988; McCarthy and Prince, 1990; Goldsmith, 1990]. It is hard to find a paper in theoretical phonology which makes no mention of prosodic structure or autosegmental tiers; phonologists are beginning to argue about whether segments exist at all.

A concern with structure follows broadly from the aim of generative phonology, which is to characterize the class of expressions which are well formed from the point of view of sound structure. Obviously this class cannot be characterized merely by providing an inventory of distinctive elements; it is also necessary to provide an account of how the elements combine in sequence. Such an account will have the general appearance of a grammar, and will therefore effectively define a structure for each allowable sequence. Even the formalism of Chomsky and Halle [1968] effectively defined structures, although these structures were very trivial compared to those proposed subsequently.

A theory of how elements combine does more than complete our theory of the phonological inventory, however. In fact, the two aspects of the theory are highly interconnected, with theoretical decisions with respect to either aspect having strong ramifications for the other. A particularly clear example of this relationship is provided by the controversy concerning whether the primitive elements (or phonemes) of tonal description are tonal targets or tonal changes. Since any single change can be specified by a sequence of two targets, all substantive contributions to this argument have rested on claims about how tonal elements behave in context, either phonologically or phonetically. As a further example, we may consider the fact that the archiphonemes have been laid to rest as a class of elements in phonological inventories. The patterns which this rather problematic construct were once held to describe are now handled by the theory of prosodic licensing, according to which prosodic nodes (especially the syllable) support different distinctive features in different positions [Goldsmith, 1990].

The interconnection of these two aspects of the theory also follows from a consideration which should be of particular concern to phoneticians. This is the fact that paradigmatic contrasts are never produced or observed in isolation. Even if an utterance consists of a single vowel *(Ahh!)*, the phonetic form of this vowel will exhibit the manifestations of its structural position, such as stress and phrase-final lengthening. Thus the interpretation of any objective evidence about paradigmatic contrasts in the phonology necessarily involves assumptions about phonological structure and its consequences. It is better for these assumptions to be laid out explicitly rather than remain implicit, because unexamined assumptions may prove to be very misleading.

The papers in the present collection all call for integrated research in phonetics and phonology. However, as a group, they only faintly reflect the single most significant recent development in phonology, which is the technical theory of structure. The experimental results presented deal either with the phonetic content of phonemes or features, or with the status of low-level rules. These issues were already apparent to readers of Chomsky and Halle [1968]. The major work on phonological structure of the last two decades are not referenced. When prosody is mentioned [as in Dressler and Moosmüller's paper, this vol.], it is treated informally rather than technically. The general implications of experimental work on prosody and suprasegmentals for the theory of sound structure are not brought out in any of the position statements. This is the case even though some of this work explicitly addresses issues such as the relational nature of distinctive features, the language-specific character of phonetic rules, and the extent to which surface phonological representations are underspecified [see Pierrehumbert and Beckman, 1988, and the literature reviewed there].

The aim of integrated research in phonology and phonetics is not merely to clarify what phonemes and features really are. Nor is it to make the job of phonologists easier by giving phonetics the responsibility for all the low-level rules. The aim is to provide a comprehensive scientific theory of language sound structure. To achieve this aim, it will be necessary for experimentalists to tackle paradigmatic contrasts and structure in an integrated fashion. There is

no theory of the phonetic content of distinctive features without a theory of the pattern of variation they exhibit in continuous speech. The interplay between articulatory and perceptual requirements on speech is exhibited in prosody as in the segmental inventory; the phonetic foundations of prosody deserve more attention. Lastly, distinctive feature theory is founded on a convergence between lexical and phonetic evidence – the dimensions of phonetic contrast which are of interest are those which distinguish words. We need to work out a similar convergence between phonetic evidence and lexical evidence concerning phonological structure and sequential constraints.

References

Beddor, P.: Predicting the structure of phonological systems. Phonetica *48:* 83–107 (1991).

Blumstein, S.: The relation between phonetics and phonology. Phonetica *48:* 108–119 (1991).

Browman, C.; Goldstein, L.: Tiers in articulatory phonology, with some implications for casual speech; in Kingston, Beckman, Papers in Laboratory Phonology I, pp. 341–376 (Cambridge University, London 1990).

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

Clements, G. N.: Akan vowel harmony: a nonlinear analysis. Harvard Studies in Phonology 2, pp. 108–177 (Indiana University Linguistics Club, Bloomington 1981).

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

Dressler, W.; Moosmüller, S.: Phonetics and phonology: a sociolinguistic framework. Phonetica *48:* 135–148 (1991).

Fourakis, M.; Port, R.: Stop epenthesis in English. J. Phonet. *14:* 197–221 (1986).

Fowler, C. A.: Comments on the contribution by Pierrehumbert and Neary. J. Phonet. *18:* 425–434 (1990).

Goldsmith, J.: Autosegmental and metrical phonology (Blackwell, Oxford 1990).

Halle, M.; Vergnaud, J.-R.: An essay on stress (MIT Press, Cambridge 1987).

Hayes, B.: The phonetics and phonology of Russian voicing assimilation; in Aronoff, Oehrle, Language sound structure, pp. 318–328 (MIT Press, Cambridge 1984).

Hayes, B.: Inalterability in CV phonology. Language *62:* 321–351 (1986).

Ito, J.: Syllable theory in prosodic phonology; Ph D diss. University of Massachusetts-Amherst (Garland, New York 1988).

Kaisse, E.; Shaw, P.: On the theory of lexical phonology. Phonol. Yb. *2:* 1–30 (1985).

Kingston, J.: Integrating articulations in the perception of vowel height. Phonetica *48:* 149–179 (1991).

Kiparsky, P.: Some consequences of lexical phonology. Phonol. Yb. *2:* 85–138 (1985).

Kohler, K. J.: The phonetics/phonology issue in the study of articulatory reduction. Phonetica *48:* 180–192 (1991).

Libermann, M. Y.; Pierrehumbert, J.: Intonational invariance under changes in pitch range and length; in Aronoff, Oehrle, Language sound structure: studies in phonology presented to Morris Halle, pp. 157–233 (MIT Press, Cambridge 1984).

Lombardi, L.: The nonlinear organization of the affricates. Nat. Lang. linguist. Theory *8:* 375–425 (1990).

Lombardi, L.: Laryngeal features and laryngeal neutralization: Ph D diss. University of Massachusetts-Amherst (1991).

McCarthy, J.: OCP effects: gemination and antigemination. Linguist. Inquiry *17:* 207–263 (1986).

McCarthy, J.: Feature geometry and dependency: a review. Phonetica *43:* 84–108 (1988).

McCarthy, J.; Prince, A.: Foot and word in prosodic morphology: the Arabic broken plural. Nat. Lang. linguist. Theory *8:* 209–238 (1990).

Maddieson, I.: Testing the universality of phonological generalizations with a phonetically specified segment database: results and limitations. Phonetica *48:* 193–206 (1991).

Marantz, A.: ReReduplication. Linguist. Inquiry *13:* 435–482 (1982).

Ohala, M.; Ohala, J. J.: Nasal epenthesis in Hindi. Phonetica *48:* 207–220 (1991).

Pierrehumbert, J.: The phonology and phonetics of English intonation; MIT Ph D diss. (Indiana University Linguistics Club, Bloomington 1980).

Pierrehumbert, J.: Phonological and phonetic representation. J. Phonet. *18:* 375–394 (1990).

Pierrehumbert, J.; Beckman, M.: Japanese tone structure. Linguist. Inquiry Monogr. Ser. 15 (MIT Press, Cambridge 1988).

Quine, W.: Two dogmas of empiricism; in From a logical point of view, pp. 20–46 (Oxford University Press, London 1964).

Selkirk, E. O.: Phonology and syntax: the relation between sound and structure (MIT Press, Cambridge 1984).

Silverman, K.; Pierrehumbert, J.: The timing of prenuclear high accents in English; in Kingston, Beckman, Papers in Laboratory Phonology I (Cambridge University Press, London 1990).

Prof. Janet Pierrehumbert
Department of Linguistics
Northwestern University
Evanston, IL 60208 (USA)