"Phonetic bases of distinctive features": Introduction

George N. Clements, Pierre Hallé

▶ **To cite this version:**

George N. Clements, Pierre Hallé. "Phonetic bases of distinctive features": Introduction. Journal of Phonetics, 2010, 1 (38), pp.3-9. 10.1016/j.wocn.2010.01.004 . halshs-00684211

HAL Id: halshs-00684211
https://shs.hal.science/halshs-00684211

Submitted on 4 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journal of Phonetics special issue

"Phonetic Bases of Distinctive Features"

Introduction

G. N. Clements and P. A. Hallé

Laboratoire de Phonétique et Phonologie (LPP)

CNRS/Paris 3 Sorbonne-nouvelle, Paris, France

1. Presentation

Distinctive features have long been involved in the study of spoken language, and in one form or another remain central to the study of phonological patterning within and across languages. However, their phonetic nature as well as their role in mental representation, speech production, and speech processing has been a matter of less agreement. Many phoneticians consider features to be too abstract for the purposes of phonetic study, and have tended to explore alternative models for representing speech (e.g., gestures, prototypes, exemplars). Psycholinguists, too, have sometimes hesitated to integrate features into their models, often preferring to work with traditional phonetic categories, segments, or syllables. The resulting breach between the representational categories of phonology on the one hand and those of the experimental speech sciences on the other has tended to increase the gap between phonology, phonetics and psycholinguistics, challenging the underpinnings of the movement to reintegrate these approaches (Laboratory Phonology). Compounding this problem is the fact that much of the experimentalist's understanding of features is still based on largely outdated theories of thirty or forty years ago, due in large part to the absence of accessible recent overviews of the subject. It would therefore seem useful to provide an up-to-date overview of the phonetic bases of distinctive feature theory as it is conceived at the present time.

The papers collected in this issue emanate, for the most part, from a conference on the theme "Phonetic Bases of Distinctive Features" held at the Carré des Sciences, Ministère Délégué de la Recherche, Paris, on July 3, 2006. This conference gathered a number of specialists from several disciplines within linguistics and the speech sciences to exchange views on the phonetic bases of distinctive features from a variety of perspectives. The larger goal was to address current issues in feature theory and to take a step toward synthesizing recent advances in order to present a current "state of the art" of the field. These brief introductory remarks will attempt to lay out the theme as it was addressed by the conference participants.[1]

2. Classical feature theories

In a widely held view whose roots stretch back to the nineteenth century and even earlier, speech sounds are defined in terms of primitive *features* corresponding to broad phonetic categories. Such features, termed "distinctive", "phonological" or "phonetic" according to the perspective adopted, are conceived of as central to the cognitive encoding of speech, which relates the variability of articulatory movements and their acoustic effects to a small number of discrete mental categories. In this view, features provide a necessary basis for understanding the structure and economy of phonological systems and provide a frame of reference for models of production and comprehension in speech communication.

Historically speaking, there have been two main trends in phonetic research on features, one emphasizing their acoustic properties and the other their articulatory properties. The first extended study of distinctive features was a short monograph entitled *Preliminaries to Speech Analysis* by Roman Jakobson, Morris Halle, and Gunnar Fant, first published in 1952 and still in print today. This collaborative effort by two phonologists (Jakobson, Halle) and an acoustician (Fant) proposed a universal set of twenty distinctive features, grouped into pairs such as *nasal* vs. *oral*, defined primarily in acoustic terms. The central hypothesis of this work was that each feature could be assigned a unique, invariant acoustic correlate (though not necessarily a unique articulatory correlate). Features could be extracted by listeners from the speech stream through the detection of their correlates and by the recognition of inter- and intrasegmental redundancies (e.g. features which never cooccur in a single segment, or which are implied by neighboring segments or the position in the word). While articulatory definitions were proposed for most features, the articulatory stage of speech was viewed as the means used to obtain each pair of acoustically contrastive effects. This point of view is summarized in the slogan "we speak to be heard in order to be understood" (p. 13).

A competing view, known as the motor theory of speech perception, was developed at about the same time at Haskins Laboratories, New Haven. Motor theory arose out of the early

finding that the acoustic patterns of synthetic speech had to be modified if an invariant phonetic percept was to be produced in different contexts (Cooper et al. 1952, Liberman et al. 1952). This work suggested that the objects of speech perception were not to be found at the acoustic level. They might, however, be sought in underlying motor processes, if it could be assumed that the acoustic variability associated with an invariant percept resulted from the temporal overlap, in different contexts, of several invariant production units. In its fullest development, motor theory held that the objects of speech perception were the intended phonetic gestures of the speaker, viewed as the elementary events of speech production and perception (Liberman & Mattingly 1985). Phonetic segments could be viewed as groups of one or more of these elementary events; for example, [b] would consist of a labial stop gesture, and [m] of that same gesture combined with a velum-lowering gesture (nasality). Phonologically, gestures could be conceived as groups of features such as *labial*, *stop*, and *nasal*, but these features were considered attributes of the gestures, not events as such. This general approach has been continued in the work of Browman & Goldstein (e.g. 1990, 1992, et seq.), who have developed a model of gestural coordination known as articulatory phonology.

The seeming radical incompatibility of these two general approaches to the study of phonological primitives has led to two competing and largely nonoverlapping traditions in feature definition. The first, represented in the work of John Ohala and others, continues to view phonological features in primarily acoustic terms. The second, following motor theory as well as the emphasis given to articulation by Chomsky & Halle (1968), gives primacy to articulatory definitions. Over the years there has been little productive interchange among the adherents of these two views. In fact, neither approach seems likely to be entirely correct. Arguing against a uniquely acoustic approach is the widely-acknowledged difficulty in finding acoustic invariants for a number of fundamental features, such as those characterizing the major places of articulation. A problem for purely articulatory approaches is raised by the

existence of articulator-independent features such as *stop* and *continuant*, which are implemented with different gestures according to the articulator employed (e.g. no invariant gesture is shared by the continuants [f], [s], and [x]). These and other problems suggest that neither a purely acoustic nor a purely articulatory account is self-sufficient.

A compromise position has sometimes been proposed in which *articulatory* features are distinguished from *acoustic* features, both having equal status. Such a view has been elaborated by Peter Ladefoged (1997). While considering the majority of features to be articulatory in nature, Ladefoged proposed six features which he considered to be best defined in acoustic terms. The feature [+sonorant], for example, is difficult to define meaningfully in terms of an articulatory invariant. While one could propose that sonorant sounds are produced with vocal cord vibration and no pressure buildup within the oral cavity, the question remains why just these two distinct articulation types are combined in a single feature. The best explanation would seem to be that they underlie the production of a group of auditorily related sounds, the class of sonorants, all of which are characterized by a periodic, well-defined formant structure.

Such an approach might seem to have the best of both worlds. However, among its current proponents, there is little agreement on what the membership of the two disjoint feature sets might be. To the extent that such theories posit novel features alongside the traditional ones, it is unclear whether they satisfy the phonological requirement of expressing the structure and content of phoneme inventories (Clements 2003, 2008) or of accounting for common phonological patterns found across languages (Mielke 2004).[2]

3. Recent developments in feature theory

We now review some of the more recent developments in feature theory. A first new impetus to the study of features has come from the *quantal theory of speech* developed by K.N. Stevens and his colleagues at MIT (Stevens 1972, 1989). A main innovation of quantal

theory is the equal status it accords to the acoustic, auditory and articulatory dimensions of spoken language. Quantal theory hypothesizes that each distinctive feature corresponds to a stable acoustic region whose auditory characteristics are not notably affected by small movements of a given articulator; larger movements of the articulator create discontinuities which define boundaries between features. Quantal theory has been applied to a certain number of features and is currently being used as a tool for the study of others, as will be illustrated by some of the following papers. (We return to a fuller discussion of quantal theory below.)

A further significant development, arising out of many detailed phonetic studies in recent years, is the increasing realization that features, and phonetic categories in general, do not necessarily have a single acoustic correlate, as was posited by Jakobson and his collaborators, but may be associated with many different cues which may be dispersed across various points in the signal. The classical example (which has perhaps received disproportionate attention in this respect) is the feature [±voice] as it is realized in English. As is well known from many studies, voiced stops are not necessarily realized with vocal fold vibration ("prevoicing"); and on the other hand, they may be associated with several cues other than vocal fold vibration, such as shorter closure duration and lengthening of the preceding vowel (see Lisker 1986 for a catalogue of cues to voicing, and Coleman 2003, Hawkins & Nguyen 2004 for evidence that the cues to voicing may be distributed over several segments). For these reasons it seems necessary to draw a strict distinction between *features*, which are located in the mind, and *cues*, which are located in the acoustic signal. From this point of view, one central task of phonetic feature definition is to identify the set of cues associated with each feature, while another is to discover how these cues are used by hearers to detect the discrete features occurring in mentally-encoded lexical representations. This somewhat more complex view of features forms the basis of a great amount of recent work in speech recognition and speech perception.[3]

A further development is the increasing recognition that a given feature is not necessarily realized with the same cue or cue set in all segment types. Thus, for example, the feature [+spread glottis] is realized as aspiration following voiceless stops, breathy voice following voiced stops, and voicelessness in sonorants, in the influential model of Halle & Stevens (1971). To at least some extent, acoustic feature definitions must be relativized to given classes of segments, even when the articulatory basis of the feature remains constant (e.g. glottal opening in the case of [+spread glottis]). This relativization may also involve differences in alignment. For example, the feature [+spread glottis] can be detected in stops only if its articulatory realization is aligned with the edge of the stop -- often the release, as in ordinary postaspirated stops, but sometimes the closure point, as in preaspirated stops. (Similar remarks can be made for other features such as [+constricted glottis] and [+strident].) The theoretical issue here is whether alignment must be specified in the definition of the feature itself, at the level at which it is coordinated with other features as in Steriade's aperture node model (1994), or at the level of gestural coordination in the sense of Browman & Goldstein's articulatory phonology.

In sum, feature theory has evolved considerably since the early work of the 1950s and 1960s, due to the development of new theoretical models on the one hand and to empirical studies that have developed our understanding of the diversity of cues that may be associated with any given feature on the other. It should be emphasized that these developments do not make features any less abstract; features remain abstract mental categories, which cannot be directly detected or measured in the signal. However, these advances have conspired to make the *study* of features more concrete, by associating them with specific articulatory states and gestures and with equally specific acoustic cues. [4]

4. Current issues in feature theory

With this background we consider the contributions to this volume, taking them up in the context of the general issues they deal with.

4.1. Biological bases of universal feature definitions

A basic goal of feature theory is to explain why languages heavily favor certain articulatory and acoustic dimensions in constructing their phoneme systems while avoiding others. The traditional explanation is that preferred contrasts maximize acoustic distinctiveness while minimizing articulatory effort. Much recent work in this direction has been carried out in the framework of dispersion theory (Liljencrants & Lindblom 1972, et seq.). However, while explaining the tendency of phoneme systems to maximize the difference between phonemes in the auditory space, dispersion theory has been less successful in accounting for the "maximal use of the available distinctive features" that typically characterizes phoneme inventories (Ohala 1980, Maddieson 1985, Schwartz et al. 1997, Clements 2003).

Quantal theory (Stevens 1972, 1989) maintains that the universal set of distinctive features can be deduced from the interactions between the articulatory parameters of speech and their acoustic effects. As mentioned above, its central claim is that for some types of articulatory parameters, there are ranges of values in which the acoustic signal is relatively stable, and that these ranges are bounded by regions in which the signal is relatively instable; the acoustic attributes of the signal within one of the stable regions define the acoustic correlates of a distinctive feature. Distinctive features are *universal* in this view, as they emerge from biological properties of the human speech production system which are essentially the same for all members of the species.

In their lead paper "Quantal theory, enhancement and overlap", Kenneth N. Stevens and S. J. Keyser offer a conceptual integration of recent research in quantal theory, enhancement

theory and gestural overlap. In their review of quantal theory, they propose that quantal relations fall into one of two general types, one arising from the aerodynamic and mechanical properties of vocal tract surfaces and the other from the acoustic filtering of vocal tract manipulation. Several examples are discussed. As they note, quantal theory seeks to explain why the inventories of distinctive features that make up the phonologies of the world's languages are what they are, but is not intended to be the principal basis of a model of speech production or lexical access; this role falls, in part, to enhancement theory, articulatory phonology, and models of speech perception and lexical access involving related notions such as landmark theory. (Enhancement theory is discussed in more detail below.)

Until recently, most work in quantal theory was based on two-dimensional models of the supralaryngeal vocal tract, including the oral, nasal and pharyngeal cavities. Current research has expanded the scope of inquiry by exploring the question of how oral side cavity resonances and subglottal resonances may create further quantal effects. The next two papers address this question. In "Subglottal resonances and distinctive features", Steven Lulich considers a quantal definition of the feature [±back] in terms of the second subglottal resonance (Sg2). This resonance is known to fall near the boundary between [-back] and [+back] vowels, and recent research has suggested that Sg2 may actually define this distinction. Lulich presents new evidence in support of this view from a study of 14 adult and 9 child speakers of American English. His primary concern is to evaluate two competing definitions of [±back]. According to the first, the boundary is defined perceptually as F3 - 3.5 bark, while according to the second, it is defined by Sg2. He found that while both definitions provide reliable boundaries between front and back vowels for speakers of all ages, Sg2 is more reliable for the youngest subjects (ages 2;2 to 9;0). In a related study of connected speech productions of an adult male speaker, he found that Sg2 forms a boundary between front and back vowels, that both Sg2 and Sg3 effectively distinguish the starting points of F2 C-to-V transitions, and that Sg3 separates the front tense vowel [i] from the front lax vowels.

Lulich suggests that such discontinuities define not only feature boundaries, but acoustic landmarks on the time dimension which may be employed in lexical access.

In "Effects of side cavities and tongue stabilization: Possible extensions of quantal theory", Kiyoshi Honda, Sayoko Takano and Hironori Takemoto consider two further factors which may contribute to creating stable regions which underlie quantal feature definitions. The first, which they term the *interdental-space effect*, involves the interdental side cavities, which are included in the oral cavity in low vowels but isolated from it by the raised tongue dorsum in nonlow vowels. As a result, somewhere in the transition between [a] and [i], there is a sudden change in the oral cavity cross-sectional area which produces discontinuities in the mid-frequency range of the second formant transitions. They suggest that this effect may provide a complementary (or perhaps alternative) account of the unstable acoustic zones that have been attributed to the coupling of the second subglottal resonance. A second factor involves the *stabilizing effect* of the co-contraction of antagonistic muscle pairs on variation in vocal tract area function. Articulatory modelling, EMG studies and MRI studies provide converging evidence that the simultaneous activation of both such pairs of muscles may help stabilize the shape of the tongue surface and thus reduce acoustic variability in certain vowels and vowel classes.

4.2. Feature theory and variation

We next take up the question of variation. General theories of sound structure such as dispersion theory and quantal theory are primarily concerned with elucidating the biological and perceptual foundations common to all languages. They are thus mainly concerned with examining invariants or universals within and across languages. However, as is well known, there is much variation in phoneme inventories and phoneme realizations within and across languages. The study of such variation constitutes a challenge for theories based on universal primitives, whether at the phonological level (features) or the articulatory or auditory level

(action theory, gestures, neural bases of audition). Most of the papers in this collection consider some of the problems raised by variation and make various proposals for treating them.

4.2.1. Features and within-language variation

Traditional feature theory combined the idea that each feature has just one correlate (or cue) with the idea that its realization was invariant in all contexts, at least in carefully articulated speech. As we know now, features may be associated with several cues, which may be present or absent depending on the segment type and the context. Moreover, important cues may be suppressed in less carefully articulated speech styles, and others may be masked by background noise or other factors that degrade the signal. A challenge for feature theory is to explain the fact that speakers can generally understand each other in spite of such impediments.

Stevens & Keyser's lead paper outlines one approach to this problem, *enhancement theory* (e.g. Stevens et al. 1986; Stevens & Keyser 1989; Diehl 1991; Kingston 1992; Keyser & Stevens 2006; Hoole & Honda 2007). In Stevens & Keyser's account, enhancement theory proposes that acoustic feature cues may have two sources. In the first place, each feature is defined by a quantal articulatory-acoustic relation and can therefore be said to be based on a defining acoustic attribute and a defining articulatory range. These defining attributes are properties of the human speech production system and are expected to be universal in language. However, additional acoustic and articulatory attributes may be added to enhance the perceptual saliency of the defining acoustic attribute. Thus, the surface representation of an utterance includes not only the feature-defining acoustic and articulatory attributes but also an array of articulatory gestures and their acoustic consequences that enhance the *perceptual* saliency of the defining attributes. These gestures are of also two general types. In one, an articulatory gesture is superimposed on the defining gesture, enhancing the defining acoustic

attribute of the feature. In the other, the supplementary acoustic attribute is separate from the defining attribute. In both cases, the enhancing attribute reinforces the perceptual cues to the feature. In Stevens & Keyser's view, the multiplicity of cues serves the function of enhancing a distinctive feature and providing further, redundant cues to its presence. They cite a number of examples in which enhancing cues may permit recovery of a distinctive feature even when its defining attribute is weakened or absent due to gestural overlap or other factors. They have also suggested (Keyser & Stevens 2006) that while defining gestures are sometimes weakened or deleted in casual speech, enhancing gestures tend to survive, preserving underlying contrasts.

Another approach to the challenge raised by variability is based on the view that not all features are lexically or phonetically specified. In an influential study, Keating (1988) examined several examples of segments that show variable realizations along given feature dimensions. She pointed out that if these features are underspecified, they will be consistent with the attributes of either value of the underspecified feature. Drawing on this approach, Lahiri and Reetz (2002) have argued for the lexical underspecification of certain features in order to explain certain behavioral asymmetries. In their proposed model of featurally underspecified lexicon (FUL), incoming speech sounds are compared online with lexically-specified features using a ternary logic of match, mismatch, and no-mismatch. Features whose cues are present in the acoustic signal do not "mismatch" underspecified segments in the lexicon, but they match or mismatch specified features. A central hypothesis is that the feature [coronal] is universally underspecified in the lexicon; as a consequence, a labial or dorsal sound detected in the signal will not mismatch an underspecified coronal feature, and can thus access it. In contrast, labial and dorsal features are specified, and so a coronal sound detected in the signal will mismatch both of these features, and cannot access them. It follows that underspecified features should exhibit a greater range of phonetic variation than specified features. Indeed, it is widely observed that plain coronal stops such as /t/ typically show a

wider range of variation within and across languages than do labial or velar stops. Their model thus accommodates certain kinds of variation, without requiring each variant to be separately listed in the lexicon or derived by a phonological process.

We have so far considered feature-based approaches to variation. It has long been recognized, however, that speech perception involves much more than the detection of acoustic feature cues. To a large extent, listeners hear what they expect to hear, regardless of what the signal contains. This indeed is a central tenet of the Chomsky/Halle view of speech perception; in their view, what a speaker "hears" is determined in large part by what the rules of phonology predict to be possible. They say,"what is perceived depends not only on the physical constitution of the signal but also on the hearer's knowledge of the language as well as on a host of extra-grammatical factors" (Chomsky & Halle 1968, 294). Most recent research on the contribution of the hearer's expectations has been carried out within the framework of psycholinguistics. It has been repeatedly shown that knowledge of the phonological system of a language, including its phonemes and its phonotactic constraints, biases the listener toward favoring certain percepts and toward disfavoring those that are inconsistent with the syllabic and phonotactic constraints of a language (Berent et al., 2007; Dupoux et al., 1999; Hallé et al., 1998; Hallé & Best, 2007). Indeed it is sometimes the case that what the hearer "hears" is in contradiction to the cues present in the signal; in such cases, linguistically-conditioned expectations may actually override the information provided by the auditory system (lexical context effects: Elman & McClelland, 1998; Ganong, 1980; phonological context effects: Massaro & Cohen, 1983; Pitt, 1998; Moreton, 2002; orthographic effects: Dijkstra et al., 1995; Hallé et al., 2000). All these factors contribute to making speech perception relatively robust in spite of variation, noise and degradation in the signal.

Not only grammatical, but extragrammatical factors play an important role in explaining the robustness of speech perception. These include:

- the hearer's (innate or acquired) knowledge of articulatory-acoustic relations, e.g. knowledge that certain sounds cannot be attributed to certain types of articulation;
- pragmatic knowledge of the world, of the topic of conversation, of what is probable vs. what is improbable in the frame of discourse;
- general laws of perception, also found in domains such as music perception and visual perception.

The last of these factors forms the main subject of Sarah Hawkins' contribution, "Phonological Features, auditory objects, and illusions". Hawkins begins by reminding the reader that features such as [+voice] or [+nasal] typically have many cues, which may be dispersed across the word. The absence of one or several cues may be compensated by the presence of others (a notion related to enhancement theory), or by recovery mechanisms that rely instead on listeners' expectations and knowledge of what is likely to occur in the speech stream at a given time. She then argues that speech perception, just like visual perception, relies on a good match between memorized experience and current sensation: when sensation meshes with expectations, listeners believe they perceive "real" linguistic objects in spite of possibly severe variation and degradation in the acoustic signal. She draws on analogies to known visual perception phenomena, including visual illusions, to suggest that the perception of distinctive features may be conditioned in part by auditory analogues of visual perception phenomena (e.g., increased perceptual salience induced by contextual contrast). Thus, Hawkins' contribution addresses the issue of the phonetic bases of feature perception in a quite novel way: the phonetic bases may be variable in nature and in location within the acoustic signal; they also are to be found in the listeners' mind; and finally, they are constrained by domain-general perceptual mechanisms.

In sum, it is increasingly recognized that what Stevens & Keyser call "defining feature cues" may be attenuated or absent in ordinary speech, yet "perceived" by listeners based on other cues, as well as on linguistically-conditioned or extragrammatical expectations. These

various factors have not yet received a full synthesis, but when they are integrated into current models of speech perception they may help explain the absence of strict invariance between distinctive features and their phonetic expression.

4.2.2. Features and cross-language variation

A second challenge to feature theory comes from the area of cross-language variation. It is well known that languages differ significantly in terms of their choice of speech sounds and in how they are realized and coarticulated with other sounds. As has often been pointed out, a given feature contrast, such as that between voiced and voiceless stops, is not necessarily realized in the same way in all languages, and indeed may show considerable variation from one language to another. If each feature were viewed as being defined in terms of the full set of its acoustic cues in any language, we would have to abandon the view that features are universal (cf. Johnson 1994). However, we would not want to fall back on the position that features may consist of any arbitrary articulatory-acoustic pairing. For example, it is unlikely, to say the least, that the feature [+voice] could be systematically realized as vocal cord vibration in one language, F2 lowering in another and high-pitched noise in a third. Feature realization is heavily constrained by acoustics and the structure of the human vocal tract.
As we have seen, enhancement theory provides one way of accommodating cross-language variation within a universalist feature theory. According to this theory, the primary cue associated with a given feature may be enhanced by other, redundant cues that are not necessarily mechanically associated with it. Notably, the attributes chosen for the purposes of enhancement may vary from one language to another. For example, in Stevens & Keyser's analysis, the feature [-voice] is enhanced by glottal opening (aspiration) in English, extending the duration of the voiceless interval and rendering voiceless stops more distinct from voiced stops; other languages, such as French, do not employ this enhancement. The general prediction is that languages may vary in their choice of non-defining attributes, and will

typically do so if these attributes contribute to enhancing the contrast made by the defining attribute.

The study of cross-language variation must of course be based on careful empirical studies of a number of languages. The basic questions include: How do languages vary in their choice of distinctive features? How do they vary in the way such features are realized? How do differences in the choice and use of features influence the way speech is perceived? These questions are directly addressed in several papers in this collection.

In their contribution "Invariant articulatory bases of the features [tense] and [spread glottis] in Korean: Stroboscopic cine-MRI data", Hyunsoon Kim, Shinji Maeda and Kiyoshi Honda provide a close examination of the phonetic bases of two distinctive features in Korean. Following a tradition launched by Chin-wu Kim (1965), these authors argue that the Korean fortis and aspirated consonants are distinguished from the lenis consonants by the feature [+tense]. Most of their discussion is concerned with defining the articulatory attributes of this feature. They show that while larynx raising is common to all Korean tense stops, tongue raising is essentially restricted to the alveolar series /t^h, t', ts^h, ts', s'/, where it may help to produce a tighter seal at the place of articulation. However, all tense consonants have longer duration than their lax counterparts, suggesting a tenser articulation at the primary place of articulation. They conclude that the feature [+tense] is articulatorily defined in terms of the simultaneous tensing of the primary articulator (lips, tongue blade, or dorsum) and the vocal folds. While other features for tense consonants have been proposed to in the literature, Kim, Honda and Maeda point out that only [+tense] has an invariant realization in all contexts. Their results are of particular interest in that the feature [+tense], unlike many other features, does not have a clear quantal definition, and (perhaps for this very reason) appears to be rarely used in languages in the absence of enhancement by other features.

Languages differ not only in their choice and use of features, but also in how the gestures that create feature contrasts are coordinated. In their paper "Perception of initial obstruent voicing is influenced by gestural organization", Catherine Best and Pierre Hallé show that differences in gestural coordination may have consequences for perception. Focusing on the perception of the voicing contrast by native speakers of American English and French, they examine three voicing contrasts involving both a coronal and a lateral constriction at word onset. These contrasts may be described as differing in their increasingly tight timing relationships between the two constrictions: loosely phased succession of dental then lateral constrictions for the Hebrew [tl]-[dl] cluster contrast, loosely synchronous constrictions for the Zulu lateral fricative contrast ([ɬ]-[ɮ]), strictly time-locked constrictions for the Tlingit laterally-released affricate contrast ([t͡ɬ]-[d͡ɮ]). Precise timing relationships of these kinds are somewhat difficult to specify in the framework of classic featural descriptions. Best and Hallé propose to use instead an Articulatory Phonology framework, in which gestures play the role of features with the additional possibility of specifying “gestural phasing relationships.” The results of the discrimination and open-response categorization tests they conducted on the three contrasts reveal increased difficulty in extracting a voicing feature value with increased tightness of gestural phasing relationship, regardless of the listeners’ native language. (Additionally, English and French listeners’ performance for Hebrew and Tlingit contrasts also reflected the difference in their native languages’ phonetic implementation of [voice] in word-initial stops.) Best and Hallé conclude that the notion of gestural timing organization in syllable onsets needs to be specified at some level of linguistic description. Their work thus suggests that feature-based linguistic descriptions must be enriched by a principled integration of timing specifications.

In their article "On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers", Keith Johnson and Molly Babel consider to what extent cross-language differences in the choice of distinctive contrasts, as well as

differences in phonological patterning, may influence speech perception. They study the perception by native speakers of Dutch vs. American English (AE) of all the 15 possible contrasts among six voiceless fricatives differing in place of articulation: [f, θ, s, ʃ, x, h], focussing on the questions: Are Dutch or American listeners insensitive to feature contrasts that are not used in their language phonemic inventory? Are Dutch listeners nevertheless sensitive to non-distinctive feature contrasts that are involved in Dutch phonological alternations (as in *poes* [s] 'cat' vs. *poesjes* [ʃ] 'kitten')? Johnson and Babel show that both Dutch and AE listeners are equally responsive to each fricative contrast at a low level of auditory perception imposed by the nature of the experimental task, suggesting they detect the acoustic cues involved in a non language-specific way at the basic auditory level. However, Dutch and AE listeners' discrimination differs for some contrasts when they are asked to rate perceptual distances. Conceivably, their performance in this task draws on higher-level linguistic representations. Johnson and Babel argue that the observed differences between Dutch and AE listeners in the treatment of fricative contrasts are partly determined by the function of the feature differences involved: whether they are contrastive, or reflect phonological alternations in the native language. Yet some contrasts escape the simple native vs. nonnative account altogether: an illustration is the AE listeners' ability to discriminate contrasts involving non-English [x]. To sum up, Johnson and Babel's work suggests that perception of a given feature (and its discrimination from other features) depends on both the feature's role in the native phonological system and a possibly universal phonetic signature.

4.3. Features in lexical access

Although lexical access is not the main focus of the present collection of papers, feature theory has often been applied to models of lexical access. Recent models include the landmark-based theory of Stevens (2000, 2005) and the FUL model proposed by Lahiri & Reetz (2002). While such models have sometimes been criticized on the grounds of their

alleged inability to account for variation, we have seen that more recent developments in feature theory have given variation full attention, and have attempted to show that variation is functionally motivated, enhancing contrasts and adding redundant information that may aid in the process of feature recognition

We have already discussed the FUL model and its implications for phonetic variation in section 4.2.1. However, Lahiri & Reetz's contribution to the present collection, "Distinctive features: Phonological underspecification in representation and processing", is primarily concerned with its application to lexical access, and we will briefly summarize their results here. Lahiri and Reetz consider the possible role of feature representations in accounting for a number of behavioral asymmetries in lexical access. They begin by reviewing a number of previous results obtained in cross-modal priming studies as well as in electrophysiological (ERP and MEG) studies. These studies offer further support for the view that [coronal] consonants and vowels are not specified in the lexicon, regardless of where they occur in the word. For example, **Homig* does not mismatch with German *Honig* ('honey'), since /n/ is underspecified for [coronal], while **Haner* does mismatch with *Hamer* ('hammer'), since /m/ is specified for [labial] and mismatches [coronal] of the surface /n/. The authors then report a novel study bearing on underspecification vs. specification of [dorsal] stem vowels in words such as *Sohn* ('son') vs. *Boot* ('boat'). They show that diminutives derived from such words always undergo umlauting, whereas plurals undergo umlauting only for words whose stem vowel is underspecified for [dorsal]. A lexical decision task with delayed priming shows that umlauted diminutive forms prime a base word only if its stem vowel is not specified for [dorsal] (e.g., *Söhnchen* primes *Sohn* but *Bötchen* does not prime *Boot*). This result is as expected under the FUL model, since the front vowel feature [coronal] of [ø] mismatches the specified back vowel feature [dorsal] of *Boot*, but does not mismatch the underspecified feature "[ ]" of the alternating vowel of *Sohn*.

5. Envoi

Space limitations have allowed us to discuss only a few of the many challenging questions and issues relating to the phonetic bases of distinctive features in this short review. We will have achieved our goal if we have whetted the reader's appetite for the varied and stimulating selection of papers that follows.

Notes

[1] We gratefully acknowledge the financial support of the Ministère Délégué de la Recherche, France, under the ACI-PROSODIE program. The full program and abstracts of the conference are available at the site http://ed268.univ-paris3.fr/lpp/phonetic-bases/program.html.

[2] See Durand (2000) for an illuminating review of the topics discussed in this section.

[3] At a time when each feature was considered to have only one cue, the terms "feature" and "cue" could be used more or less interchangeably in acoustic studies. This terminological practice was inadvertently carried over into work explicitly recognizing the multiplicity of cues, as in the title of the classic paper cited above as Lisker (1986), which could more accurately be named "A catalogue of acoustic *cues*...".

[4] See Hall (2007) for an up-to-date list of the most commonly-used distinctive features.

References

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, *104*, 591-630.

Browman, C., & Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, *18*, 299-320.

Browman, C., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155-180.

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

Clements, G.N. (2003). Feature economy in sound systems. *Phonology*, *20*, 287-333.

Clements, G.N. (2008). The role of features in speech sound inventories. In E. Raimy & C. Cairns (Eds.), *Contemporary Views on Architecture and Representations in Phonological Theory* (pp. xx-yy). Cambridge, MA: MIT Press.

Coleman, J. (2003). Discovering the acoustic correlates of phonological contrasts. *Phonetica*, *31*, 351-372.

Cooper, F., Delattre, P., Liberman, A., Borst, J., & Gerstman, L. (1952). Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America*, *24*, 597-606.

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

Dijkstra, T., Fieuws, S., & Roelofs, A. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology, 49***,** 264–271.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1568-1578.

Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes*, *16*, 491-505.

Durand, J. (2000). Les traits phonologiques et le débat articulation/audition. In P. Busuttil (Ed.) *Points d'interrogation : Phonétique et phonologie de l'anglais* (pp. 56-70). Pau: Presses Universitaires de Pau.

Elman, J., & McClelland, J. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, 27, 143–65.

Ganong, W. (1980). Phonetic categorization in auditory perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 110-125.

Hall, T. A. (2007). Segmental features. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology* (pp. 311-334). Cambridge: Cambridge University Press.

Halle, M., & Stevens, K. (1971). A note on laryngeal features. *Quarterly Progress Report of the Research Laboratory of Electronics* (MIT), *101*, 198-213. [Reprinted in: Halle, M. (2002). *From Memory to Speech and Back: Papers on Phonetics and Phonology 1954-2002* (pp. 45-61). Berlin: Mouton de Gruyter.]

Hallé, P., & Best, C. (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *Journal of the Acoustical Society of America*, *121*, 2899-2914.

Hallé, P., Chéreau, C., & Segui, J. (2000). Where is the /b/ in "absurde" [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language*, *43*, 618-639.

Hallé, P., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance, 24*, 592-608.

Hawkins, S., & Nguyen, N. (2004). Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English. *Journal of Phonetics*, *32*, 199-231.

Hoole, P., & Honda, K. (to appear). Automaticity vs. feature-enhancement in the control of segmental F0. In G.N. Clements & R. Ridouane (Eds.), *Where do features come from? The nature and sources of phonological primitives.* Elsevier.

Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to Speech Analysis: the Distinctive Features and their Correlates*. Cambridge, MA: MIT Press.

Johnson, K. (1994). Phonetic arbitrariness and the input problem: Comments on Coleman's paper. In P. Keating (Ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III* (pp. 325-330). Cambridge: Cambridge University Press.

Keating, P. (1988). Underspecification in phonetics. *Phonology*, *5*, 275-292.

Keyser, S., & Stevens, K. (2006). Enhancement and overlap in the speech chain. *Language*, *82*, 33-63

Kim, C.-W. (1965). On the autonomy of the tensity feature in stop classification. *Word*, *21*, 339-359.

Kingston, J. (1992). The phonetics and phonology of perceptually motivated articulatory covariation. *Language and Speech*, *35*, 99-113.

Ladefoged, P. (1997). Linguistic Phonetic Descriptions. In W. Hardcastle & J. Laver (Eds.), *The Handbook of Phonetic Sciences* (pp. 589-618). Oxford: Blackwell.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven, N. Werner, & T. Rietveld (Eds.), *Laboratory Phonology 7* (pp. 637-676). Berlin: Mouton.

Liberman, A., Delattre, P., & Cooper, F. (1952). The role of selected stimulus-variables in the perception of the unvoiced stop consonants. *American Journal of Psychology*, *65*, 497-516.

Liberman, A., & Mattingly, I. (1985). The motor theory of speech perception revisited. *Cognition*, *21*, 1-36.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language*, *48*, 839-862.

Lisker, L. (1986). 'Voicing' in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech*, *19*, 3–11.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

Massaro, D., & Cohen, M. (1983). Phonological context in speech perception. *Perception and Psychophysics, 34*, 338-348.

Mielke, J. (2004). *The emergence of distinctive features*. Unpublished doctoral dissertation, The Ohio State University. Columbus, OH.

Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, *84*, 55-71.

Ohala, J. (1980). Moderator's introduction to the Symposium on phonetic universals in phonological systems and their explanation. In *Proceedings of the 9th International Congress of Phonetic Sciences* (vol. 3, pp. 181-185). University of Copenhagen: Institute of Phonetics.

Pitt, M. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception and Psychophysics, 60*, 941-951.

Schwartz, J.-L., Boë, L.-J., Vallée, N., & Abry, C. (1997). The dispersion-focalization theory of vowel systems. *Journal of Phonetics*, *25*, 255-286.

Steriade, D. (1994). Complex onsets as single segments: the Mazateco pattern. In J. Cole & C. Kisseberth (Eds.), *Perspectives in Phonology* (pp. 203-291). Stanford, CA: CSLI Publications.

Stevens, K. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In E. David Jr. & P. Denes (Eds.), *Human Communication: a Unified View* (pp. 51-66). New York: McGraw-Hill.

Stevens, K. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*, 3-46.

Stevens, K. (2000). Diverse acoustic cues at consonantal landmarks. *Phonetica*, *57*, 139-151.

Stevens, K. (2005). Features in speech perception and lexical access. In D. Pisoni & R. Remez (Eds.), *Handbook of Speech Perception* (pp. 125-155). Cambridge, MA: Blackwell Publishers.

Stevens, K., & Keyser, S. (1989). Primary features and their enhancement in consonants. *Language*, *65*.1, 81-106.

Stevens, K., Keyser, S., & Kawasaki, H. **(**1986). Toward a phonetic and phonological investigation of redundant features. In J. Perkell & D.Klatt (Eds.), *Symposium on Invariance and Variability of Speech Processes* (pp. 426-463). Hillsdale, NJ: Lawrence Erlbaum.