PHONETIC INTERPRETATION IN GENERATIVE GRAMMAR

Geoff Lindsey and John Harris

Abstract

This paper presents an alternative to the widely held view that the phonological component of a generative grammar is an engine which converts cognitive lexical representations into ever more physical phonetic ones. Our view is that phonological representations are uniformly cognitive, while phonetic events are wholly specified in performance theory. Nevertheless, elementary phonological units map onto perceptual representations of the acoustic signal in a relatively direct manner. We illustrate this mapping with a phonological and spectrographic analysis of a number of segmental decomposition processes in English.

0 Introduction

> General phonetics has been neglected for so long that agreement on even the most elementary propositions of phonetic theory cannot be taken for granted at present (Chomsky & Halle 1968: 298).

More than twenty years after the publication of *The Sound Pattern of English*, it seems that little has changed: the domain of enquiry of general phonetics remains as ill-understood as ever. So much so that it is currently possible for linguists to assign to 'phonetics' any low-level phenomena relevant to speech, such as facts of anatomy and aerodynamics, and for speech technologists to assign to 'phonetics' any high-level phenomena relevant to speech, such as phonotactic constraints and morphological structure. There does not even appear to be a consensus as to whether phonetics is a part of linguistics, either conceptually or in the real world of academic administration.

One view stemming from SPE, and assumed with varying degrees of explicitness in the literature, is that phonology is an engine which transforms cognitive lexical representations into ever more physical phonetic ones. Our purpose in this paper is to motivate an alternative view in which phonological representations, specified by the phonological component of the grammar faculty, are always purely cognitive. Phonetics by contrast is not 'modular', either in a cognitive or in a physical sense, but is rather a cover term. Phonetic events are wholly specified in performance theory (namely that part which is concerned with the specification of speech sounds). Performance theory subsumes a theory of grammatical competence, of which phonology is a component.

We begin in 1.1 by pointing out some of the disagreements that exist in the literature regarding the nature of the phonetics-phonology interface and in 1.2 we present two possible views of the interface. We espouse one of these and take issue with a recent manifestation of the other (Bromberger & Halle 1989). In 1.3, we argue for phonological representations which are most naturally mapped into auditory rather than articulatory descriptions. In the second half of the paper, our aim is to demonstrate the relatively direct manner in which elementary phonological units are mapped onto perceptual representations of the acoustic signal. In 2.1 and 2.2, we outline an analysis of certain lenition processes in English which is based on the notion of segmental decomposition. On the basis of these phonological representations, we discuss with reference to spectrographic data how phonetic characterizations can be assigned to the posited elements.

1 Phonetics and grammar

1.1 Where is phonetics? According to SPE, 'phonetics is concerned with grammatically determined aspects of the signal' (p. 294). Some of the tenets of SPE have remained intact over the past twenty years while others have been substantially revised or replaced. Little further work has been done during this time to develop or criticize the views of SPE on the status of phonetics in generative grammar, but it would be a mistake to conclude from this that these views constitute a coherent, exhaustive account that has been widely adopted. A close reading of SPE and a comparison of it with accounts in the subsequent literature, such as those in established textbooks on generative phonology, reveal certain discrepancies and indeterminacies.

For example, *phonetic* transcriptions in the terms of SPE itself 'consistently disregard many overt physical properties of speech. Among these are *phonetic* effects...such as the voice pitch and quality of the speaker and also such socially determined aspects of speech as the normal rate of utterance' (p. 295; our italics). These 'phonetic' effects cannot belong to grammar and so must be outside the domain of 'phonetics' as SPE defines it on the preceding page.

In Sommerstein's (1977) account of generative phonology, systematic phonetic representations have 'a degree of "narrowness" such that, at the very least, any two sounds that are distinguished in any human language are differently represented' (p. 115). This conception is less 'narrow' than that of SPE itself, which acknowledges a [suction] phonetic feature on the grounds that languages like Yoruba produce labiovelars with suction while languages like Late do not, 'regardless of the fact that apparently in no language are there contrasting pairs of utterances that differ solely in this feature' (p. 297-8). For SPE, universally non-contrastive signal information that is systematically present in a language must be included in phonetics unless it is present *universally*.

Hyman (1975: 171) gives the following examples: 'a [k] will tend to be somewhat fronted before [i] and a vowel will tend to be somewhat nasalized before a nasal consonant. These universal tendencies are, as such, not part of the phonologies of individual languages, but rather belong to the realm of universal phonetics'. It is not clear whether Hyman's realm of universal phonetics should lie within grammar. For Kenstowicz and Kisseberth (1979: 179), a phonetic representation 'includes all grammatically determined aspects of pronunciation (as opposed to aspects of pronunciation determined by physiology, for instance)'. To the extent that the tendencies mentioned by Hyman are determined by physiology, then, they do not belong in phonetic representations, while somehow being a part of phonetics.

The SPE statement that 'phonetics is concerned with grammatically determined aspects of the signal' defines the relevance of phonetic theory to a theory of competence while making it clear that its data (the signal) is to be found in performance (in contrast to phonological and syntactic theory, data for which may be obtained by introspection). Thus a phonetic transcription is 'a representation of what the speaker of a language *takes to be* the phonetic properties of an utterance, given his hypothesis as to its surface structure and his knowledge of the rules of the phonological component' (p. 294; our italics). Phonetic information is viewed as a part of speakers' knowledge of their language. So the goal of phonetic theory is to characterize this knowledge rather than to characterize the signal itself, although the theory must be constructed with reference to the signal.

Some of the confusion regarding the status of phonetics seems to have resulted from this implied dual role which comprises (i) the setting up of phonetic *representations*, which are considered to belong at the surface 'end' of the phonological component of the grammar, and (ii) the mapping of these representations into physical sounds, which of course lies outside the grammar.

1.2 The phonetics-phonology interface: two views. If phonetics is viewed at least in part as characterizing speakers' knowledge of speech sounds, the question of the definition of phonology and of phonological representations arises. The phonological component in generative grammar relates surface representations to underlying ones, the latter being stored in lexical memory. The former are sometimes called phonetic representations and the latter sometimes phonological representations, but SPE characterizes both of them in terms of phonetic features (although even here there is an occasional reference to the features in their underlying manifestations as 'phonological features'). There are differing views on the correct distribution of speech sound information amongst the three levels: underlying (lexical-level) representation, surface (sentence-level) representation and signal (utterance-level) representation. We will now try to define for the purposes of discussion two views of the way in which this information may be distributed. In spite of the confusion just discussed, we suspect that the following view coincides with assumptions that are implicit in much of the literature.

(1) *Distinctiveness-based view*

(a) Underlying representations contain the minimal information necessary to distinguish one lexical item from another is a given language.
(b) Surface representations contain the minimal information necessary to distinguish one speech sound from another across languages (see the comments regarding Sommerstein above).
(c) The signal contains all other information.

The view we adopt here is rather different:

(2) *Competence/performance-based view*

There are two levels of constraints on speech signals:
(a) Cognitive, grammatical constraints, common to all members of a speech community and independent of any specific utterance.
(b) Non-grammatical constraints of various subtypes:
(i) physical constraints common to all normal humans; and
(ii) physical constraints peculiar to each individual.

We define the constraints on speech signals of type (2a) as 'phonology', which is thus a subpart of the theory of competence. Constraints of type (2b) belong in a theory of performance, and not in competence. However, performance subsumes competence; and for the part of performance theory which concerns speech sounds, and which therefore incorporates all of (2a) and (2b) above, we use the term 'phonetics'. Thus a phonological interpretation is analogous to a strictly semantic interpretation of sentence meaning, and a phonetic interpretation (subsuming its phonological interpretation) is analogous to a full pragmatic interpretation of utterance meaning. Note that (1) and (2) express orthogonal points of view: their truth values are independent. The phonological constraints of view (2) *might*, for example, be such as to maintain intra- and inter-linguistic distinctiveness with minimal information, but adherence to (2) does not require this.

The second part of this paper will propose phonetic mappings for several basic phonological units and discusses the consequences for the signal of the ways in which these units may combine and de-combine to form surface representations. The constraints on combination, and the units themselves, are phonological and therefore cognitive; but it will be shown that the cognitive representations map rather directly into signal effects, and indeed that these cognitive representations can be inferred rather directly from signal patterns which we assume are extracted by the human perceptual mechanisms from the signal's physical quantities.

The most recent expression of the (1)-type view by one of SPE's authors is tto be found in Bromberger and Halle 1989 (hereafter B&H). The picture sketched by B&H of the distribution of phonological and phonetic information seems to us incoherent as a competence-theoretic account of phonology, in that it is dynamic and physical. According to B&H (p. 53), 'Since underlying phonological representations of words are stored in speakers' permanent memory, whereas phonetic surface representations are generated only when a word figures in an actual utterance, there is a clear and theoretically significant sense in which underlying representations are prior to surface representations'. We are unable to make sense of this assertion other than in the context of a performance model. Generativity, we assume, describes that property of grammatical knowledge which ensures that all and only well-formed surface structures are assigned an interpretation. It does not seem to us possible to assert that this property of grammar only obtains on the occasions of specific utterances by speakers.

B&H further consider that underlying and surface representations should encode information in different ways on the basis of the functions which they serve: memory and articulation/audition, respectively. We assume that phonological representations are cognitive and by definition serve cognitive functions. We do not believe that their encoding should be designed to serve the functions of articulation and audition, which are performance functions. Syntactic information, for example, may *possibly* be encoded in such a way as to serve certain discourse functions, but this is not obviously true, and likewise we do not wish to make assumptions regarding the synchronic relevance of non-grammatical functions to the form of cognitive representations.

Nonetheless, we agree with B&H that, whereas 'syntax is concerned with the relations among representations that encode different types of information requiring different types of notation, phonology is concerned with the relationship between representations that encode the same type of information' (p. 53). In ways which will become clear in Section 2, we consider that the units manipulated by the principles of the phonological component are phonetically interpretable at all points, and in the same way.

1.3 Articulation and audition. Finally in this section, we wish to point out the incompatibility of B&H's assertion that surface representations serve the functions of 'articulation *and audition*' (our italics) with the particular form of phonological representation which they subsequently present. According to their footnote 5 (p. 54), 'For present purposes we restrict attention exclusively to the articulatory aspect of language and ignore the auditory interpretive system'. It is not clear to us how, for any other than their present purposes, their analyses could be couched in auditory terms, other than by extra transformations of an *ad hoc* and uneconomic nature. Crucially, their phonological surface representations 'must serve as input to our articulatory machinery' (p. 53). In accordance with the performance-oriented dynamicism and functionalism of their view, and following the feature-geometric approach originating in the work of Clements (1985) and Sagey (1986), the structure of their phonological representations models the physiology of speech production. Thus their representation of [bel] in terms of geometrically-organized features 'specifies the vocal tract gymnastics necessary for uttering the word *bell*' (p. 54).

There is no reason for phonological representations to recapitulate the performance mechanisms of speech production, *viz.* the articulators. Phonological representations are couched rather in terms of the cognitive categories by reference to which the child makes sense of the speech sounds which it hears. There may be an evolutionary link between the structure of these categories and the structure of the hearing and speaking mechanisms, but this link is not necessarily relevant to the characterization of grammatical knowledge. According to B&H (p. 57-8), on the other hand, the difference between underlying and surface representations

> is ultimately a consequence of the fact that like all physical systems the individual articulators are subject to inertia and their movements are influenced by their earlier positions and movements and by simultaneous movements and positions of other articulators. Although these contextual effects have their origin in mechanical factors, they achieve certain articulatory optimizations that are brought about in different ways by different languages. They are thus not mere effects of the physics or physiology involved.

Phonological processes, then, are 'articulatory optimizations' of physical effects, and phonological representations encode this fact. We, however, wish to exclude performance constraints from consideration in positing phonological representations. Consider the perspective of Dell (1980: 35):

> The only regularities that can be ignored [in phonology] are those that are found in all languages and attributed to certain constraints imposed by the structure of the human articulatory or perceptual system. For example, it is useless to note the difference between the nasal bilabial consonant at the end of *in* in *in between* and the labiodental nasal which often replaces it in *in fact*. There is no rule that systematically distinguishes these two types of sounds in any language, and labiodental nasals only regularly replace bilabial nasals when the latter immediately precede an [f] or a [v].

However, it is exactly this sort of (often phonologically irrelevant) phenomenon that the articulation model of B&H is designed to capture (and, indeed, predicts to be of primary phonological relevance).

Further, there is evidence to suggest, in moving beyond phonology and into phonetics, that the cognitive categories should be mappable in the first instance not into articulations but into sound patterns (*sic*). We assume, following SPE (pp. 296-7), that the aim of phonological analysis is to set up representations and mechanisms such that each generalization captured reflects 'a "regularity" of the sort that a child will use as a way of organizing the data he is confronted with in the course of language acquisition'. The language learner has auditory access to a full range of adult speech sounds, but can receive by vision only a partial impression of adult speech articulations (in addition, of course, to the 'data' of the child's own immature productive apparatus). Accordingly, perception precedes production in the process of language acquisition (see Edwards 1974). Likewise, the congenitally deaf never acquire the ability to produce speech naturally, even with intensive explicit training, while the congenitally dumb are apparently able to acquire the normal skills of speech perception (see Fourcin 1975).

Such facts follow rather naturally from a framework in which the phonological categories of universal grammar constitute expectations regarding the structures to be inferred from auditorily processed acoustic signals. We suggest that the study of articulatory gestures has a status in linguistics not unlike the study of visual representations of spoken language, i.e. writing systems. Hearing, but neither speech nor sight, is a necessary condition (under otherwise normal circumstances) for phonological competence to grow in the mind. As Andersen (1974: 42-3) points out in the context of a discussion of phonological change,

> articulatory phonetics, which is so important for the training of linguists, is quite irrelevant for an understanding of abductive innovations in phonology.

> It is apparently possible for some linguists to believe that people communicate primarily by means of vocal gestures, rather than by auditory symbols -- I am referring to the incredible 'motor theory of speech perception.' But it is obvious that the language learner who acquires a phonology, in the first place, bases his analysis on auditory sense data and, in the second place, in his gradual acquisition of the motor skills involved in speech production, monitors the results of his articulatory motions by ear. This means that as a historical linguist, with a theory of language change which focusses attention on the real processes in which linguistic innovations arise, such as language acquisition, one would like one's colleagues in experimental phonetics to invest more work on the relation between acoustic phonetics and speech perception. It is sad that the efforts initiated by Roman Jakobson to provide acoustic definitions of distinctive features were not continued and further developed in this direction. Chomsky and Halle's (1968) supposed revision of Jakobson's distinctive feature framework -- with which in fact it has nothing in common -- was a giant step in the wrong direction. In anachronistically reintroducing articulatory definitions, the authors ignore the lesson, repeatedly emphasized by their teacher, that the auditory experience is the communicative process actually shared by both speaker and addressee...

We stress that the grammar manipulates signal-based patterns rather than sound waves, or for that matter vocal tract (or cochlear) gymnastics. We do not wish to replace one physically based dogma with another. This is clearly a question not of mutually exclusive alternatives but of primacy or 'basicness'. Chomsky (1986) rightly objects to a view of the grammar as solely a cognitive *input* system (in the sense of Fodor 1983), 'if only because it is used in speaking and thought ... the input and output systems must each access a fixed system of knowledge' (p. 14). Our contention is that this knowledge is innately couched in a form which responds *primarily* to sound data rather than to visual and proprioceptive evidence of articulations by others and by oneself. Obviously, a comprehensive theory of language use will ultimately need to characterise the phonological output system, i.e. map phonological entities into (some representation of) the speech production physiology. It may be that this sub-theory will have some of the characteristics of feature geometry theory and/or of gestural 'phonology' (Browman and Goldstein 1989), but note that even this sub-theory will have to capture the necessity for normal speech production of *feedback* from primary, sound-based representations.

Keating (1985: 16) has pointed out the extent to which articulatory or acoustic orientation in the phonetic study of phonological categories has followed technological innovation. It is slightly ironic that today, as the technological sophistication of acoustic analysis increases so extraordinarily in response to the pressing concerns of automated speech synthesis and recognition, many phonologists should adopt the articulation-based assumptions of feature geometry theory and look for phonetic reassurance in the study of physiological gestures. A phonological orientation exclusively towards physiological mappings minimises contact not only with speech synthesis and recognition research but also with such areas of cognitive science as modularity, parsing and learnability theory -- in fact all areas concerning interaction between individuals and the data to which they are exposed.

2 Phonetic interpretation of phonological elements

2.0 Our aim in this section is to demonstrate the relatively direct manner in which elementary phonological units are mapped onto the acoustic signal. The main source of evidence that we will exploit in seeking to identify these units is one that is familiar to phonologists: the behaviour of segments under phonological processing. The reasoning involved here is that it is frequently only when a segment moves that it reveals its stripes. The processes we focus on all involve the lenition of *t* in various types of English.

2.1 *t*-lenition in English. In the following discussion, our phonological and phonetic data illustrate three lenition processes: glottalling (as in Glasgow, London, etc.), tapping (as in most dialects of North American English) and spirantization (as in Liverpool, Dublin, etc.). The specific conditions under which these processes occur coincide to a large extent. Informally, we may refer to the primary contexts as foot-internal and domain-final position:

(3)	Foot-internal	Domain-final
	city	*fit*
	letter	*get*
	Peter	*feet*
	photographic	*boat*

We will have nothing further to say about the characterization of these conditions, since our main concern is the nature of the processes themselves. (For a government-based account of the relevant conditions, see Harris & Kaye, in press, and Harris, in press.)

Glottalling of *t* results in debuccalization to [ʔ] and in London occurs both foot-internally and domain-finally. In tapping systems such as New York City, *t*-lenition produces an alveolar tap in foot-internal position. Domain-finally, however, the tap only occurs if a word-initial vowel follows. Otherwise we find an unreleased preglottalized *t*:

(4)	Tap [ɾ]	Unreleased [ʔt˺]	
	Prevocalic	Prepausal	Preconsonantal
	fit us	*fit*	*fit me*
	late again	*late*	*late mark*

In Liverpool, spirantized *t* is typically realized as an alveolar slit fricative with non-strident sibilance. It is potentially distinct from *s* (*letter* ≠ *lesser*), although these sounds are variably merged in some speakers. Lenition here extends to *k* (e.g. [bæx] *back*). The conditions under which *t*-spirantization occurs are more or less identical to those operating in the case of London glottalling. In function words, the process can proceed as far as a complete loss of oral gesture, as in [nɒh] *not*, [bʊh] *but*, [æh] *at*. Non-lenited plosive [tʰ] and lenited [s], [ʔ] and [ɾ], are illustrated in the wide band spectrograms and speech waveforms in Figs. 1-4, where the relevant forms occur foot-internally in *daughter* .

2.2 Lenition as segmental decomposition. The analysis of these data that we assume here is essentially that presented in Harris, in press. There it is argued that all lenition events should be directly represented in terms of segmental decomposition, i.e. as the loss of material from the internal structure of a segment. This goal is only achievable within a theory in which phonological oppositions are expressed privatively, as in Dependency Phonology (Anderson & Ewen 1987), Government Phonology (Kaye, Lowenstamm & Vergnaud 1985), Particle Phonology (Schane 1984) and the 'extended' dependency approach of van der Hulst (1989). Within an orthodox binary-valued feature framework in which contrasts are uniformly represented in equipollent terms, lenition processes can only ever be characterized as the replacement of one feature-value by another. The set of possible replacement operations is impressively large but unfortunately contains a significant number of predicted process-types that are simply not attested in natural language.

In contrast, the Government approach we adopt here incorporates a radically pared-down theory of phonological activity. Only two types of phonological activity are countenanced: composition, in which material from one segment spreads onto a neighbouring segment, and decomposition, in which material is lost from the internal structure of a segment. When linked to a strictly privative system in which oppositions are uniformly expressed in terms of univalent elements, this impoverished model of phonological processing generates an extremely small set of possible decomposition operations. The number of decomposition processes which a given segment can potentially undergo is logically limited by the number of elements of which it is composed.

Terms such as lenition or weakening, when applied to consonantal reduction phenomena, make implicit appeal to the notion of phonological strength. According to the usual textbook definitions of strength, a segment X is deemed stronger than segment Y if X passes through a Y stage on its way to deletion (Hyman 1975: 165). Thus a typical lenition sequence is [t > s > h > Ø], which involves spirantization followed by debuccalization, followed by deletion. This is taken as evidence of a hierarchy of strength relations, in which [t] is the strongest and [h] the weakest segment. (The preferences expressed in such hierarchies are identified on a cross-linguistic basis by methods of comparative and internal reconstruction. Within a given system, two or more stages in a sequence may co-exist as alternants. There is no implication that lenition along a particular trajectory inevitably leads to deletion in every case.)

If the assumption that reduction processes are to be represented as element depletion is correct, it follows that lenition along a particular strength hierarchy involves a progressive decrease in the elementary complexity of a segment. It then makes sense to suppose that the least complex segment, the one occupying the stage immediately prior to zero, contains only one element. From the perspective of a framework in which phonological elements are held to have independent phonetic interpretation, this observation is highly significant: pre-deletion stages in lenition chains allow us actually to 'hear' individual elements. Besides [h], other 'primitive' segments which show up as pre-deletion targets in lenition are [ʔ] (which like [h] results from debuccalization) and [w, j, ɾ, ɯ] (all the result of vocalization). Given our line of argumentation, each of these should be the phonetic instantiation of a particular element. And, indeed, three of these segments are independently motivated as the manifestation of elements in vocalic systems, namely [w] = U, [j] = I and [ɯ] = the maximally unmarked 'cold' vowel (see Kaye, Lowenstamm & Vergnaud 1985). Our immediate concern is to present some of the phonological evidence which indicates that the other primitive segments, i.e. [ɾ, ʔ, h], also correspond to particular elements, symbolized here as R, ʔ and h.

In Government Phonology, it is proposed that the phonetic identity of each element is composed of one marked or salient property and a number of unmarked or background properties. Compound segments are formed by element fusion, an operation that involves headed pairs of elements. In a fused expression, the non-head contributes only its salient property; all other properties are contributed by the head. The autonomous phonetic identity of an element thus only manifests itself when it occurs as the head of a simplex segment.

The lenition facts indicate that the independent phonetic identities of the elements ʔ, h and R are respectively glottal stop, glottal fricative and coronal tap. Following Kaye, Lowenstamm & Vergnaud (1988) and Harris (in press), we will assume that the salient properties of these elements, characterized phonetically at this stage merely with informal labels, are as follows:

(5)	Element	Salient property	Independent manifestation
	ʔ	occlusion	[ʔ]
	h	noise	[h]
	R	coronality	[ɾ]

Below we attempt to provide more detailed sound pattern specifications.

Fusion of ʔ and R produces a coronal stop. We will assume that a true (i.e. released) plosive is distinguished from an unreleased stop by the additional presence of h. The combination of h and R in a bi-elementary structure defines a coronal fricative.

The analysis of *t*-lenition as element depletion can now be formalized as follows:

(6)	(a)	(b)	(c)	(d)	(e)
	x	x	x	x	x
	|	|	|	|	|
	ʔ	ʔ	|	|	|
	|		|	|	|
	R		R	R	|
	|			|	|
	h			h	h

(a) plosive *t*
(b) glottalling (debuccalization to [ʔ])
(c) tapping ([ɾ])
(d) spirantization to [s]
(e) debuccalization to [h]

2.3 Signal mappings of elements. Consideration of the spectrograms and waveforms in Figs. 1-4 suggests that this conception of lenition as material loss is mirrored in the acoustic facts. Taking the aspirated plosive [tʰ] of *daughter* in standard southern British pronunciation, we can list the acoustic properties of this utterance which are not due purely to the surrounding vowels as follows:

(7) (a) Formant transition: here a sudden rise in F2 between [ɔ] and [t].
(b) Interval of significantly reduced amplitude: a brief period with little energy on either spectrogram or waveform.
(c) Aperiodic energy: the release burst and subsequent noise corresponding to the aspirated quality.

Comparison of this with the lenited forms shows that each lacks one or two of these properties. The tap [ɾ] clearly has the rapid F2 transition but not the aperiodic energy. The glottal stop lacks a rapid F2 transition and aperiodic energy, and here has a clear silent interval (though this is not always the case - see below). Conversely the spirantized reflex of *t* clearly has both the local F2 rise and aperiodic energy, but lacks a silent interval.

The status of the reduced amplitude interval in [ɾ] is less obvious. We predict absence of the ʔ element, though acoustically there is a very brief amplitude reduction here. [ɾ] seems to be accompanied by fairly regular vocal fold vibration. Arguably what counts as 'significant' in a reduced-amplitude interval is either reduction to silence (as in [tʰ], and in this [ʔ], but not in [ɾ]) or an interval of substantial duration (as in [ʔ] but not [ɾ]). Even in the presence of the ʔ element, then, the absolute peak-to-peak amplitude envelope could thus remain quite high; impressionistically, however, there is a marked drop in loudness (and it is clearly desirable to test the perceptual effects of various manipulations of the amplitude envelope). It is worth pointing out that property (7c) corresponds to the 'abrupt' or [-continuant] feature of Jakobson & Halle (1956: 41-2), defined acoustically as 'silence (at least in the frequency range above the vocal cord vibration) followed and/or preceded by a spread of energy over a wide frequency region, either as a burst or as a rapid transition of vowel formants'. This characterization is equally problematic in the face of the spectrographic detail of [ʔ] and [ɾ]

On the basis of these observations we suggest that sound patterns (7a), (7b) and (7c) are associated (though not uniquely) with the salient properties of phonological elements R, ʔ and h respectively. This yields a rather direct mapping between the phonological account of lenition above and the acoustic facts. Of course, (7a), (7b) and (7c) are not full specifications of the acoustic mappings of these elements. R is not simply mapped onto a particular transition but rather the locus frequency towards which the formant tends; further, formant transitions such as that visible in the [s] of Fig. 2 are often absent before voiceless fricatives, in which case 'coronality' is cued rather by the frequency range of the fricative. It is commonly and cross-linguistically the case that [ʔ] is characterized not by simple cessation of voicing throughout but by an abrupt drop in Fo and a marked change to creaky phonation type (Lindsey, Haruna & Hayward 1990). The elements are thus mapped into sets of patterns (not all of which must be embodied in the signal).

Against the possible objection that a set of superficially disparate auditory properties is less elegant than reassuringly unitary articulatory labels such as 'comonal', it must be stressed that these properties are mapped onto a single phonological entity and are not individually accessible by the grammar.* What is significant is that the phonological loss of material proposed here for lenition predicts the loss of precisely the correct sets of acoustic properties. Each element is a 'superpattern' (to which we may give an arbitrary label such as 'coronality') subsuming one or more 'subpatterns' (such as high or low band of noise, high or low F2 locus), each of which in turn must be mappable onto a range of physical values. Only the superpatterns are manipulable by the phonology.

References

Andersen, H. 1974. Towards a typology of change: bifurcating changes and binary relations. *Proceedings of the First International Conference on Historical Linguistics*..

Anderson, J. & C. Ewen. 1987. *Dependency Phonology*. Cambridge: Cambridge University Press.

*This point is made by Kaye, Lowenstamm & Vergnaud when they offer a definition of vocalic elements that is couched in terms of orthodox SPE features: 'features may not be accessed directly or manipulated in any way [by the theory]. Their role is rather a secondary one serving as an instrument of phonetic interpretation of phonological elements' (1985: 306).

Bromberger, S. & M. Halle. 1989. Why phonology is different. *Linguistic Inquiry* 20. 51-70.
Browman, C. P. & L. Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6. 201-251.
Chomsky, N. 1986. *Knowledge of language: its nature, origin, and use*. New York: Praeger.
Chomsky, N. & M. Halle 1968. *The sound pattern of English*. New York: Harper & Row.
Clements, G. N. 1985. The geometry of phonological features. *Phonology* 2. 223-50.
Dell, F. 1980. *Generative phonology and French phonology*. Trans. C. Cullen. Cambridge: Cambridge University Press.
Edwards, M. L. 1974. Perception and production in child phonology: the testing of four hypotheses. *Journal of Child Language* 1. 205-219.
Fodor, J. A. 1983. *The modularity of mind*. Cambridge, MA: MIT Press.
Fourcin, A. J. 1975. Speech perception in the absence of speech productive ability. In N O'Connor (ed.), *Language, cognitive deficits, and retardation*. London: Butterworth.
Harris, J. In press. Segmental complexity and phonological government. *Phonology* 7:2.
Harris, J. & J. Kaye. In press. A tale of two cities: London glottalling and New York City tapping. To appear in *The Linguistic Review*.
Hulst, H. van der. 1989. Atoms of segmental structure: components, gestures and dependency. *Phonology* 6. 253-84.
Hyman, L.M. 1975. *Phonology: theory and analysis*. New York: Holt, Rinehard and Winston.
Jakobson, R. & M. Halle. 1956. *Fundamentals of language*. The Hague: Mouton.
Kaye, J., J. Lowenstamm & J-R. Vergnaud. 1985. The internal structure of phonological elements: a theory of charm and government. *Phonology* 2. 303-26.
Kaye, J., J. Lowenstamm & J-R. Vergnaud. 1988. Konstituentenstruktur and Rektion in der Phonologie. In M. Prinzhorn (ed.), *Phonologie*, Wiesbaden: Westdeutscher Verlag. To appear as 'Constituent structure and government in phonology' in *Phonology* 7:2.
Keating, P. 1985. The phonology-phonetics interface. *UCLA Working Papers in Phonetics* 62.
Kenstowicz, M. & C. Kisseberth. 1979. *Generative phonology*. New York: Academic Press.
Lindsey, G., A. Haruna, A. & K. Hayward. 1990. Hausa glottalic consonants: a laryngographic study. M.s.
Sagey, E. 1986. The representation of features and relations in non-linear phonology. PhD diss. Cambridge, MA: MIT.
Schane, S.S. 1984. The fundamentals of Particle Phonology. *Phonology* 1. 129-55.
Sommerstein, A.H. 1977. *Modern phonology*. London: Edward Arnold.

Figure 1. Speech waveform and wide band spectrogram of *daughter* with intervocalic plosive, adult male speaker. Divisions of vertical axis: 500 Hz. Time axis: 500 ms.

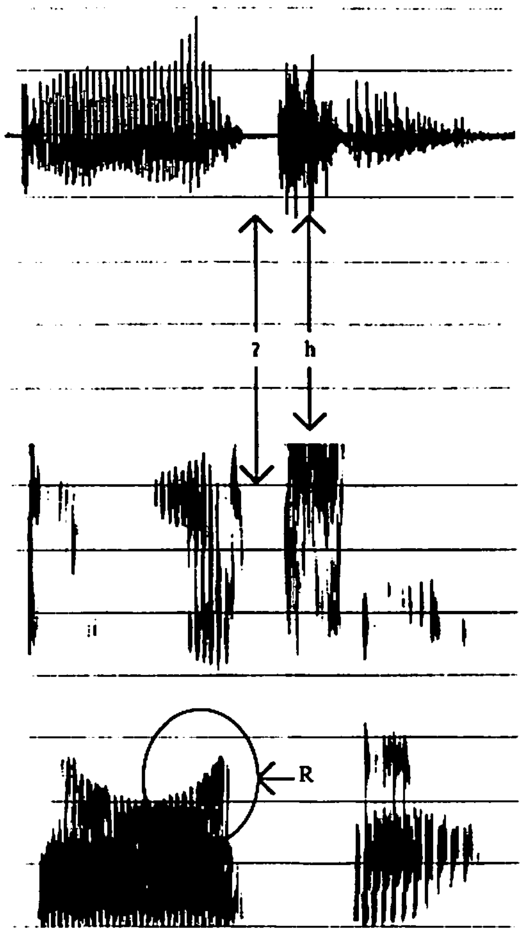

Figure 2. Speech waveform and wide band spectrogram of *daughter* with intervocalic fricative, adult male speaker. Divisions of vertical axis: 500 Hz. Time axis: 500 ms.

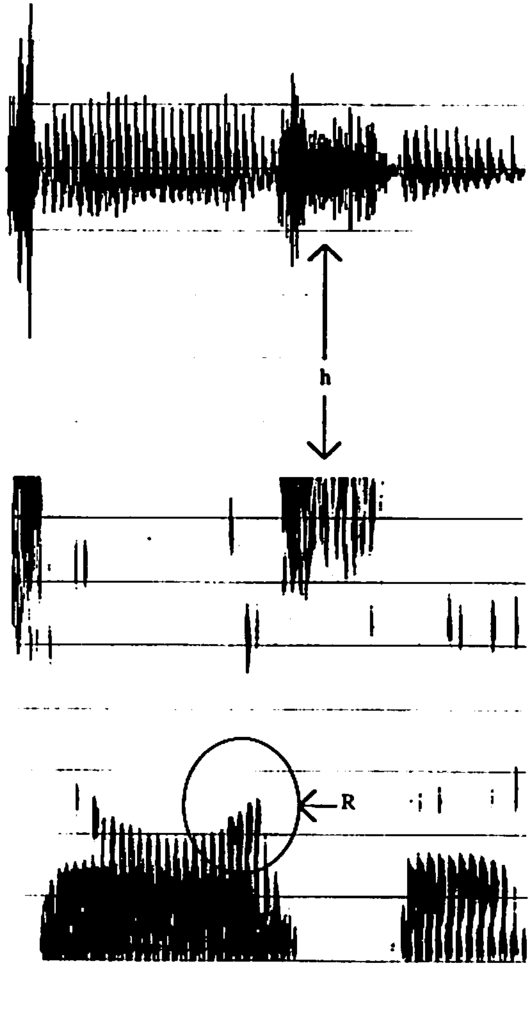

Figure 3. Speech waveform and wide band spectrogram of *daughter* with intervocalic glottal stop, adult male speaker. Divisions of vertical axis: 500 Hz. Time axis: 500 ms.

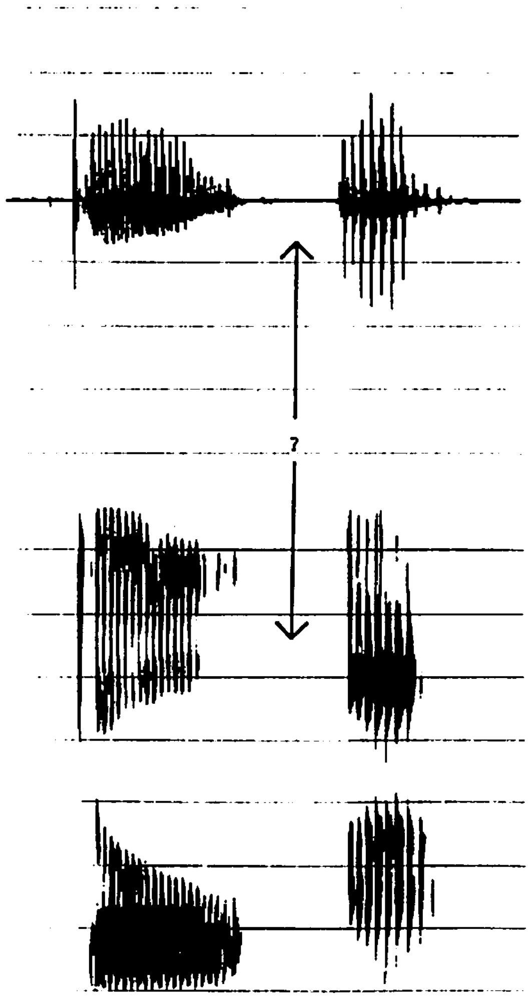

Figure 4. Speech waveform and wide band spectrogram of *daughter* with intervocalic tap, adult male speaker. Divisions of vertical axis: 500 Hz. Time axis: 500 ms.