Speech, Hearing and Language: work in progress

Volume 13

Lenition degrades information: consonant allophony in Ibibio

John HARRIS and Eno-Abasi URUA

Department of Phonetics and Linguistics
UNIVERSITY COLLEGE LONDON

Lenition degrades information: consonant allophony in Ibibio

John HARRIS & Eno-Abasi URUA

Abstract
Consonantal lenition degrades information in the speech signal and should be understood as having an analogous impact on phonological representations. The point is illustrated by a phonological and instrumental study of Ibibio (Lower Cross, Nigeria). The more general claim is that informational asymmetries in the speech signal are matched by informational asymmetries in phonology.

1. Introduction

There is a clear sense in which consonantal lenition can be said to degrade information in the speech signal.[1] A lenited consonant projects fewer cues onto the acoustic signal than an unlenited congener. In this paper, we suggest that lenition should be understood as having an analogous impact on phonological information — the code in terms of which the sound shape of morphemes is compiled. More generally, informational asymmetries in the speech signal are matched by informational asymmetries in phonology.

Making the case for this approach rests on the assumption that there exists a much tighter relation between phonology and the speech signal than is allowed for by the heavy articulatory bias of orthodox feature theory. In this respect, the treatment of lenition presented here differs sharply from recent accounts which have tended to dwell on its articulatory aspects, particularly on the notion that it is driven by a need to minimise articulatory effort (see for example Kirchner 1998, in press, Lavoie 2000). To be sure, the evident similarities between articulatory undershoot in speech, presumably related to effort, and entrenched lenition in phonology can hardly be coincidental. Parallels of this sort reflect the more general point that phonology is shaped in non-trivial ways by functional forces, including those with a basis in articulation. We may wish to view these forces as being embodied in synchronic constraints which actively determine the output of phonological grammars (Flemming 1995, Kirchner 1998, Boersma 1999). Or we may see them as purely historical, influencing the evolution of phonological systems over time (see for example Hyman, in press, Hale & Reiss 2000, and the work of Ohala (1992, 1995). Either way, we should not allow a concern with the articulatory aspects of lenition to divert attention from the following fundamental question: what impact does it have on the capacity of phonological representations to convey linguistic information?

Informational details of this sort cannot be directly captured by standard feature theory, a point we discuss in §2. Instead what is called for is a model that is not shackled to a narrowly phonemic conception of contrast and that posits a close connection between segmental categories and properties of the speech signal. A model meeting these requirements is presented in §4

[1]Eno Urua is at the University of Uyo, Nigeria. We are grateful to Mark Huckvale for his invaluable advice and help in customising the speech analysis software used in this study. Thanks also to Gordon Hunter for help with the statistical analyses and to Andy Faulkner for statistical advice and helpful comments on an early draft of the paper.

Lenition is no respecter of phonemic status: it robs a consonant of information regardless of whether some or all of that information would traditionally qualify as contrastive. This point is well illustrated in Ibibio, the language spotlighted in this paper.[2] When non-initial in the stem, oral stops in Ibibio are subject to three processes which result in the suppression of acoustic cues associated with stem-initial plosives: devoicing, loss of release, and spirantisation (the conversion of a stop into a continuant). Under standard assumptions, only one of these — devoicing — would count as having contrastive import in Ibibio: a voicing distinction that is supported stem-initially is neutralised stem-finally. The others would count as merely allophonic. But all three processes have in common the fact that they reduce the amount of information supported outside stem-initial position. Devoicing suppresses periodicity in obstruents; loss of release suppresses the noise burst associated with plosion; and spirantisation suppresses the sustained interval of radically reduced amplitude associated with stop closure. The overall picture of Ibibio that emerges is an asymmetric one, presented in §5, in which consonants in stem-initial position are informationally richer than those in other contexts.

Spirantisation and loss of burst release present a descriptive challenge which impacts on the whole question of what counts as phonologically informative. Impressionistic observation of these phenomena in different languages indicates that their effects are often variable and phonetically continuous, a point discussed in §3. Ibibio appears to be typical in this respect. In §6, we present a quantitative study of four Ibibio speakers which sets out to determine the informational potential of two acoustic cues targeted by lenition — the reduced-amplitude interval associated with stop closure and the aperiodic energy associated with continuous frication or plosive release. Specifically, we investigate the extent to which leniting and non-leniting contexts can be reliably distinguished on the basis of these properties. More generally, the study sets out to bring experimental evidence to bear on the hypothesised relation between cue-based segmental categories and the speech signal.

2. The auditory basis of phonological information

2.1 Phonological information and signal information

Put yourself in the shoes of a listener seeking to identify an intervocalic *p*. The speech signal will contain a number of cues to the identity of the consonant, including the following: an abrupt change in amplitude, a noise burst, rapid formant transitions, and perturbations in fundamental frequency. The value of these cues derives from a simple perceptual principle: change is more salient than stability. Speech, like any communication system, involves modulations of a carrier signal (cf. Traunmüller 1994). The greater the magnitude of a modulation, the more easily detectable it is (Ohala & Kawasaki 1997: 14). The salience of the cues in question stems from the fact that they impose spectral discontinuities on the carrier signal — landmarks which stand out against the background of periodicity associated with the surrounding vowels.

Compare this situation with one in which the listener is called on to identify an intervocalic *w*. In this case, the most reliable type of cue is provided by the formant

[2] Ibibio, spoken mainly in Akwa Ibom State, Nigeria, belongs to the Lower Cross family of Delta Cross (Benue-Congo) languages.

transitions between the glide and the surrounding vowels. The smoothness of the transitions in this case makes this particular cue rather less salient than the corresponding rapidly changing cue associated with *p*. Similarly, any amplitude change that might occur during the glide is likely to be much less marked and abrupt than in the case of the plosive.

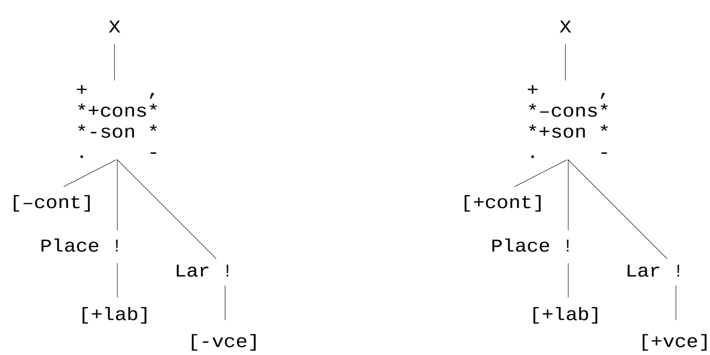
If we think of the slate of signal cues associated with a segment as a reflection of its information-bearing capacity, it is clear that, in the same intervocalic context, a plosive projects more information than a glide:

(1)

	V*p*V	V*w*V
Formant transitions	✓	✓
Abrupt amplitude drop	✓	✗
Noise burst	✓	✗
F_0 discontinuity	✓	✗

Signal information of the sort exemplified in (1) maps to phonological information — the categories in terms of which the phonetic form of morphemes is represented. In current phonological theory, the most influential model of categorical representation continues to be some version of that established by the *Sound Pattern of English* (Chomsky & Halle 1968). If phonological information is conceived of as being constructed out of SPE-style features, the relation to signal information can hardly be considered direct. To see this, compare (1) with the corresponding geometric representations containing the SPE-derived features in (2).

(2) *p* *w*



Of the various feature values represented here, some can be considered to map relatively directly to particular signal cues: [–continuant] corresponds to the amplitude drop in *p*, and [+labial] corresponds to the formant transitions in both *p* and *w*.

In general, however, it would be fair to say that phonetic-informational asymmetries of the sort displayed in (1) are not transparently reflected in representations of the type illustrated in (2). There are three main reasons for this. First, the bivalency of standard features grants the same status to the presence of some item of information as to its absence. Second, certain perceptually salient properties of the speech signal are denied independent featural status on the grounds that they are contextually predictable and thus phonemically non-contrastive, the plosive-release example mentioned earlier being a case in point. Third, the explicitly articulatory orientation of standard features renders them ill-suited to the task of establishing a close relation between the phonological code and the speech signal. Let us consider each of these points in more detail.

2.2 Bivalency

Standard feature specification is founded on the traditionally equipollent notion of phonemic contrast. This encourages the false expectation that, for every piece of information borne by a given feature value, there will be a corresponding piece borne by its complement. As far as signal information is concerned, this turns out to be incorrect in at least two ways. First, the complement value may not correspond to a uniquely identifiable property in the signal. For example, while [+labial] relates to a unifiable set of spectral properties, the same cannot be said of [–labial]. There is no obvious set of spectral attributes that unites, say, alveolars, velars and pharyngeals. Second, certain background properties associated with the carrier signal are granted equal representational status with informationally salient cues. For example, periodicity in glides is treated to a whole array of specifications in (2) — [–consonantal, +sonorant, +continuant, +voice].

This particular shortcoming can of course be alleviated by ditching bivalent specification in favour of privativeness. For example, the informational asymmetry inherent in contrasts based on the amplitude drop in plosives can be expressed by means of a monovalent feature such as [stop] or by some geometric manner node.

2.3 Monosystematicity

The phonemic conception of contrast incorporates the assumption that the distinctive resources of a phonological system should be reduced to a single context-free inventory. This too can be shown to render opaque the relation between the phonological code and signal information.

A sound property qualifying as phonemically contrastive typically meets two criteria. First, it enters into a relation of local contrast: that is, it participates in a paradigmatic distinction holding in a particular phonological context. Second, it can be correlated with a sound property that is locally contrastive in some other context. In English, for example, a place contrast holding locally in word-initial position (as in **pin, tin, kin**) can be correlated with a local place contrast in word-final position (**nip, nit, nick**). The second criterion gives rise to a monosystemic view of contrast which serves the purpose of alphabetic economy but which downplays the informative potential of certain sound properties. Some locally contrastive properties which cannot readily be cross-contextualised are known to be perceptually highly salient. In English, the aspiration and the noise burst associated with plosives are of this type (Abramson & Lisker 1970). Under a standard monosystemic account, these properties are deemed non-contrastive on the grounds that they are predictably tied to particular contexts. In feature terms, they are treated as the redundant reflexes of specifications which can be cross-contextualised — [–voice] in the case of aspiration, [–son, –cont] in the case of release burst.

The high cue potential of certain supposedly redundant properties rests to a large extent on their very predictability. Aspiration in English is not only paradigmatically informative, acting as the most robust local cue to the 'voice' identity of plosives, but it is also syntagmatically informative to the extent that it adheres to the onset of a stressed syllable and thus demarcates the left edge of a foot.

Any approach to segmental categorisation that attempts to give full due to the syntagmatic informativeness of certain sound properties has to break with the monosystemic tradition of mainstream feature theory.

2.4 Articulatory features

There is another, rather obvious reason for the mismatch between standard features and cue information. In contemplating the linguistic informativeness of signal cues, we are taking the perspective of the listener. SPE-type features are defined in the first instance from the perspective of the speaker. This articulatory bias is even more heavily accentuated in more recent geometric offshoots (Clements 1985, McCarthy 1992, Clements & Hume 1995) which, in recapitulating details of vocal-tract anatomy, clearly resemble certain models of speech production — especially articulatory phonology (Browman & Goldstein 1989) and mini-tracts theory (Stevens & Keyser 1994). The explicitness of this articulatory orientation, it has to be acknowledged, partly immunises standard feature theory against the criticism that it fails to connect in any immediate way with acoustic cue information: it was not specifically designed to do so.

Nevertheless, it is clear that any feature theory has to make the connection with the speech signal at some point. In the case of the standard model, this necessitates reference to some supplementary mechanism which translates articulatory representations into auditory-acoustic specifications. There is no obvious way of implementing this notion other than by embracing the view that the listener perceives linguistic messages in terms of the talker's articulatory movements — either by internally synthesising the intended movements (the motor theory of speech perception, cf. Liberman & Mattingly 1985), or by literally hearing articulatory gestures rather than speech signals (the direct-realist theory of speech perception, cf. Fowler 1986). The idea that the events perceived in speech are vocal-tract manoeuvres invites comparison with deaf sign, where gestures undeniably do constitute the objects of perception (in this case visual). There are many who remain unconvinced of the need to depart from what surely counts as the null hypothesis — that the listener perceives proximate acoustic signals rather than the distal articulations that produce them (see for example Ohala 1986, Lieberman & Blumstein 1988: 147ff., Klatt 1989, Moore 1989: 273ff.).

The alternative view — that phonological representations are fundamentally auditory-acoustic in nature — is the one with the deeper roots in the history of phonological theory, associated with the writings of Saussure, Sapir, Jakobson and others. Embracing it allows for the link with signal information to be made in a much more direct manner. Articulations, on this view, are external to the phonological code; they have no specification other than in terms of the continuously varying motor control mechanisms that speakers activate in order to achieve auditory-acoustic targets — the 'auditory theory of speech production' (Ladefoged, DeClerk, Lindau & Papcun 1972).

Auditory-acoustic features have long figured in work directly concerned with speech perception (see Liberman & Blumstein 1988 (188 ff.) for a review of the relevant literature). They have also undergone something of a revival in recent phonological theory (see for example Flemming 1995, Steriade 1997, Boersma 1998, and the contributions to Hume & Johnson, in press). Especially where this has occurred under the auspices of 'phonetically driven' OT, the justification most often appealed to embodies two quite independent claims: (i) many phonological phenomena are motivated by pressures emanating from the auditory-acoustic domain, and (ii) these pressures are directly incorporated in the grammar in the form of active constraints which must have access to auditory-acoustic features. The first of these claims is surely beyond controversy. The second, however, is open to question on the grounds that the phenomena in question can be adequately explained by reference to general theories of speech perception, language change, and language acquisition, without having to seek additional, grammar-internal motivations (again see the work of Ohala (1992)).

Whatever the pros and cons of this controversy might be, it should not be allowed to obscure what seems to us to be the most fundamental motivation for asuming that phonological features are intrinsically auditory-perceptual — their function as conveyors of linguistic information. Phonology can be understood as the code by which morphemes are translated into sound and vice-versa. The point at which the hearer's and the speaker's experience of this translation intersect is the speech signal.[3] But the informational potential of the signal is only realised through being processed by the auditory-perceptual system. A reasonable first hypothesis consistent with this observation is the one Sapir, Jakobson and others arrived at: the communicative contract between speaker and hearer is underwritten by a shared auditory imagery, and it is in terms of this imagery that the phonological code is specified. Not only is the imagery what the listener makes reference to in decoding a linguistic message, but it also provides the speaker with the target to be aimed at when encoding a message and monitoring its transmission. Any deviation from this position represents a research hypothesis that should not be adopted hastily. This applies particularly to the notion that SPE feature theory has long taken for granted, that the phonological code should be specified in articulatory terms. Articulatory manoeuvres deliver linguistic messages, but it is not obvious that they are in and of themselves any more informative than the hand movements of a sign-writer or the electrical activity of automated traffic lights.[4]

It would not matter where in the speech chain we chose to define the phonological code, if there existed some simple isomorphic relation between adjacent links that would allow us to translate directly between them. However, we know this not to be the case. The relation between the different links displays a clear unidirectionality, which is particularly marked at the production end of the chain: the same acoustic effect can be achieved by different articulatory means.

One of the main arguments mounted in favour of the motor theory of speech perception is that acoustic cues to a given segment can be quite diverse and contextually variable and can only be unified by reference to the articulatory gestures that must have produced them. In response to the point that articulations are themselves also highly variable, the theory proposes that different instances of the same gesture display a topological unity that sets them apart from instances of other gestures (Liberman & Mattingly 1985: 22). The sheer diversity of articulatory

[3]Even the motor theory acknowledges that the acoustic signal serves as source of 'information' about articulatory gestures (Liberman & Mattingly 1985: 12).

[4]Visual perception of articulations can of course play a role in speech comprehension (see Schwartz, Robert-Ribes & Escudier 1998 for a review of the relevant literature). However, certain rather obvious facts underline the point that vision is ancilliary to audition in speech. For example, speech is entirely intelligible without visual input, and the congenitally blind acquire normal phonology.

manoeuvres that can collaborate to produce particular acoustic effects makes this position difficult to maintain.

Consider the example of vowel quality, the production of which is known to involve a high degree of variability across different languages, across different speakers, and even within the speech of a single speaker (Ladefoged *et al.* 1972, Lindblom, Lubker & Gay 1979). As a result, no unique inversion is possible from vocalic speech signals to vocal tract area function (see for example Atal, Chang, Matthews & Tukey 1978). A particular vowel quality is not uniquely definable in terms of specific tongue or lip postures (such as 'height' or 'backness') but is determined rather by the overall shape of the vocal airway. This makes it often impossible to attribute shared topology to articulations which converge on a particular quality but which are located in quite different regions of the vocal tract. To take one specific instance of this: a 'flattening' of the sound spectrum (a downward shift in a set of formants) can be achieved by either lip rounding or pharyngeal contraction or both (Jakobson, Fant & Halle 1952: 31).

The same point can be made with respect to consonantal quality. Obstruent voicing provides a good example. It is a well known fact that, in order to maintain the transglottal airflow necessary for vocal-fold vibration, some active compensatory measure has to be taken to reduce the build-up of supralaryngeal air. This can be achieved by a variety of means, for example by widening the pharynx or venting the naso-pharyngeal port (see for example Halle & Stevens 1971, Rothenberg 1968). The production of a single acoustic property — periodicity in obstruents — can thus involve harnessing an array of articulators — larynx, pharynx, velum — which are topologically quite diverse.

The unidirectionality of the relation between different stages of the speech chain supports the conclusion that phonological information is distilled in the auditory-perceptual domain. As Jakobson, Fant & Halle put it, the 'closer we are in our investigation to the destination of the message (i.e. its perception by the receiver), the more accurately can we gage the information conveyed by its sound shape' (1952: 12).

Returning to the V*p*V and V*w*V examples above, we can imagine the signal effects listed in (1) as being produced by a variety of vocal tract shapes. However, it is reasonable to suppose that the effects map in a fairly direct manner to particular auditory-perceptual responses. If we take seriously the proposition that the link between signal information and phonological information is reasonably transparent, then each of the cues in (1) can be considered to correspond to a piece of the featural code. Before pursuing this point in more detail, we need to say more about the nature of the relation between the segments in (1).

3. Lenition: the categorisation problem

Within a given language, *p* and *w* may enter into a relation traditionally known as lenition. That is, they may occur as alternants or distributional variants, typically reflecting a historical process whereby a plosive is weakened to a glide in certain phonological contexts. The particular type of lenition which converts a noncontinuant into a continuant is widely attested in the world's languages. Probably the best studied case is Spanish (James Harris 1969, 1984, Mascaró 1984 and the references there).

(For further examples, see Kirchner's (1998) fully referenced language survey.) This type of lenition is also firmly established in Ibibio.

The examples in give a preliminary idea of the form lenition takes in Ibibio. Root-final oral consonants show up as stops when utterance-final or preconsonantal but as continuants when prevocalic.

(3)	*dɨ́p*	'hide'	*dɨ́βé*	'hide oneself'
	bèt	'shut'	*bèɾé*	'be shut'
	fʌ́k	'cover'	*fʌ́ɣɔ́*	'cover oneself'

The description and analysis of lenition in Ibibio raise a number of difficulties that will be familiar to anyone who has come across this phenomenon in other languages.

Firstly, it is not always easy to characterise the lenited reflexes in terms of traditional impressionistic articulatory labels. The broad transcriptions in (3) are those widely applied to this sort of lenition and are typical in that they gloss over gradient variability in the degree of stricture involved. In Ibibio, this variability is to be found both dialectally and within the speech of individual speakers. Connell's (1991) detailed phonetic description of the language refers variously to 'tapped approximants', 'tapped fricatives', 'tapped stops', 'approximant-like quality', 'weak, unstable articulation' (1991: 65ff.).

In the case of labial and dorsal reflexes, the IPA categorisation underlying the use of the symbols *β* and *ɣ* is potentially quite misleading. The extension of the two-way voice classification across all obstruents encourages the expectation that languages will exhibit this distinction in fricatives no less than in plosives. In fact, voiced fricatives are known to be highly marked (Ladefoged & Maddieson 1996: 176ff.). Indeed, in many cases, segments reported as voiced fricatives turn out on closer inspection to be frictionless continuants — in other words, not obstruents at all. There is a good articulatory reason for this: vocal fold vibration inhibits airflow, thus reducing the potential for air turbulence at the point of stricture (Ohala 1983).

Impressionistically, it is difficult to determine whether individual variants transcribable as *β* or *ɣ* in Ibibio qualify as full-blown fricatives; they certainly do not exhibit the high level of frication noise associated with prototypical strident fricatives. In this respect, the situation is similar to that prevailing in other languages where the term SPIRANTISATION is traditionally applied to a lenition process that more often than not gives rise to frictionless continuants rather than spirants (cf. Kirchner 1998: ch 4, Lavoie 2000). This is reportedly the case in Spanish, for example (Cole, Hualde & Iskarous 1998). This suggests that another traditional term referring to lenition — VOCALISATION — may often be more appropriate.

It is probably true to say that all languages possessing genuine voiced fricatives (French and Polish, for example) also have homorganic voiceless counterparts (cf. Ladefoged & Maddison 1996: ch 5). In contrast, where so-called spirantisation gives rise to a voiced continuant series, this is not necessarily matched by an existing homorganic voiceless set. In Ibibio, there is no voiceless series corresponding to *β/ɾ/ɣ*.

These matters of classification might have been dismissed as trivial transcription worries, were it not for the substantive implications they have for an information-based approach to phonological representation. The main issue to be taken up here is this: how much information projected by an unlenited segment is retained by a spirantised or vocalised congener?

Another issue, of relevance to output-oriented theory in general, concerns the derivational thinking implicit in the labels traditionally attached to different types of lenition. Describing a segment as spirantised or vocalised begs the question of what it is a spirantised or vocalised reflex of.

The classically derivational answer is that a lenited segment issues from a 'phonemically' or 'underlyingly' or 'canonically' unlenited segment, typically a stop. This solution is particularly favoured if a stop congener can be found outside the lenition site (see for example James Harris's (1969) treatment of spirantisation in Spanish). A well-known case is Danish, where a medial contrast between lenis stops and a vocalised series is matched to an initial contrast between fortis and lenis stops respectively (Jakobson, Fant & Halle 1952, Davidsen-Nielsen 1978). Establishing relations of this sort becomes notoriously tricky when lenition neutralises a contrast that is maintained in other positions. For example, is the tapped reflex of coronal stops in English to be associated with canonical *t* or *d*? In Ibibio, as suggested by the alternations in , intervocalic *ß/ɾ/ɣ* can evidently be related to utterance-final *p/t/k*. But the picture is clouded by the fact that, on the basis of impressionistic observation, the final stops can be described as typically unreleased (as in many dialects of English). From an information-centred viewpoint, this effect also qualifies as lenition, on the grounds that it results in the suppression of a cue (a noise burst) which is present in initial plosives. The latter support a voice distinction which is suspended in both of the lenition contexts illustrated in (exemplification to follow). Ibibio thus throws up a more general version of the phonemic conundrum raised by tapping in English: are *ß/ɾ/ɣ* to be related to the voiced or voiceless series of initial stops? (Reasons of phonetic similarity might have favoured association with the voiced series, were it not for the fact that Ibibio lacks a *g* to match *ɣ*.)

Neutralisation undermines any attempt to subsume segments in different contexts under a common canonical form (Twaddell 1935, Trubetzkoy 1939) — or, to use rather more modern parlance, to derive different surface segments from a single underlier. It would not be entirely unfair to conclude that the practice is motivated primarily by considerations of alphabetic hygiene. In any event, the thinking behind it is inimical to an approach which seeks to establish a direct link between signal cues and phonological categories.

To return to the inter-segment relation illustrated in : lenition potentially diminishes the amount of information a segment projects onto the speech signal. For example, a vocalised segment lacks a selection of the cues that are present in an unvocalised congener. Lexical access to some canonical stop form of a lenited segment would thus require the listener to reconstruct — to hallucinate — features for which there is no direct evidence in the speech signal. This in itself is not entirely far-fetched; for example, there is evidence that listeners are able to restore phonetic information that has been experimentally excised from the signal or masked by non-speech noise (Warren 1970, Warren & Obusek 1971). However, there is a significantly simpler view of signal-to-lexicon mapping which dispenses with the notion of canonical segments altogether. It is one in which lexical entries contain fragmentary feature representations rather than fully specified phoneme-like entities (Lahiri & Marslen-Wilson 1991). There is no place in this model for abstract pre-lenition segments: the low degree of signal information presented by a historically weakened segment maps directly to a low degree of categorial representation in the lexicon.

On this view, it is potentially misleading to speak of the speech signal as containing 'cues to segment x'. The cues map to individual features rather than to segments *per se* (cf. Stevens & Blumstein 1981). Moreover, it is not necessary to assume, as for example Lahiri & Marslen-Wilson (1991) do, that an individual feature is only phonetically interpretable once it is harnessed to a full span of other features that combine to define some canonical segment. Giving up this assumption represents a significant departure from standard feature theory, in which 'fragmentary' applied to phonological representation implies 'underspecified', which in turn implies 'phonetically uninterpretable until missing feature values have been filled in' (Halle 1959, Archangeli & Pulleyblank 1994). Forging a transparent link between signal cues and the phonological code requires a feature model in which segments containing minimal information are no less phonetically interpretable than those that are informationally much better endowed. What such a model might look like is the subject of the next section.

4. Cue-based elements

4.1 Modelling informational asymmetries

Any proposal to specify segmental categories in exclusively auditory-acoustic terms inevitably owes a hefty intellectual debt to Jakobson, Fant & Halle (1952). However, if we are to capture the informational imbalances inherent in segmental distinctions, there are significant respects, already touched on above, in which a cue-based model must part company with the Jakobsonian and SPE traditions.

First, by abandoning bivalency in favour of privativeness, we avoid having to grant categorial status to properties of the speech signal that have no more than a background carrier function.

Second, by departing from a monosystemic conception of contrast, we are in a position to grant categorial status to salient signal cues which, through being tied to particular contexts, are deemed non-distinctive in phonemic-based approaches. A polysystemic approach, in the Firthian tradition, gives full recognition to the syntagmatically informative nature of such properties (see Robins 1970 for discussion and references).

Third, we give up the assumption that the phonetic interpretability of a given feature specification is contingent on its being integrated into a segment containing other specifications. Instead we assume that each feature is endowed with the ability to be expressed in isolation. This has the desirable result of divorcing the informational complexity of a segment from its phonetic interpretability: a segment such as *w*, associated with minimal cue information, is no less interpretable than a segment such as *p*, with higher informational yield.

With reference to the *p*–*w* example, we adopt the specific proposal that each of the cues in corresponds to a phonological ELEMENT (Lindsey & Harris 1990, Brockhaus, Ingleby & Chalfont 1994, Harris & Lindsey 1995, 2000, Williams 1998a, 1998b, Brockhaus & Ingleby in press, Harris in press).[5] Abstracting away from the laryngeal dimension for the moment, let us focus on the three element-cue correspondences defined in (4)a.

(4)		SIGNAL CUE	ELEMENT
	(a)	Abrupt and sustained amplitude drop	[edge]
		Noise burst	[noise]
		'Labial' spectral pattern	[rump]
	(b)	Labial plosive	[rump, edge, noise]
		Labial approximant	[rump]

As set out in (4)b, a labial plosive contains three elements. Its manner properties are defined by [edge] and [noise]. [edge] corresponds to the silent interval produced by stop closure. [noise] corresponds to the burst transient and subsequent aperiodic energy produced by turbulent airflow on the release of closure. The same element also characterises the continuous noise associated with fricatives. The place specification of the segment in b is defined by [rump]. This manifests itself in the vicinity of the stop as a marked skewing of acoustic energy to the lower end of the spectrum — the 'diffuse-falling' spectral pattern identified by Blumstein & Stevens (1981). On its own, [rump] corresponds to the set of formant transitions associated with a labial approximant (Harris & Lindsey 2000).

The element [edge] is equivalent to Jakobson, Fant & Halle's (1952) [abrupt] and SPE's [–continuant]. The element [noise], on the other hand, has no precedent in orthodox feature theory, in which aperiodic energy receives two quite distinct treatments. When associated with continuous frication, it is subsumed under [+continuant] alongside the periodic energy associated with vocoids. When associated with the burst transient accompanying plosive release, it has no independent featural representation whatsoever. The element [rump] — elsewhere designated by the mnemonic [U] — can perhaps be viewed as an amalgam of [grave] and [flat] in the Jakobson, Fant & Halle (1952) feature system (Anderson & Ewen 1987). It is only very roughy equivalent to [+labial] or [+round], since bottom-heavy spectral tilt can be achieved by articulatory means other than lip protrusion.

The representations in b illustrate how the element model allows the informational load borne by a segment to be treated independently of its phonetic interpretability. Phonologically, there is a clear informational asymmetry between a plosive, bearing three elements, and a glide, with one. In spite of its impoverished specification, the glide enjoys full phonetic interpretability. It is 'primitive' in the sense that it embodies a single element, [rump], interpreted in isolation. Its ability to be made phonetically manifest is not contingent on anything equivalent to the filling-in of redundant values required by standard feature theory. From the perspective of element theory, operations of this sort are not only unnecessary but also undesirable, for the reason that they would obscure the very informational imbalances we are attempting to capture.

[5] The proposal arises out of the general tradition represented in the work of, *inter alios*, Anderson & Jones (1974), Schane (1984), Kaye, Lowenstamm & Vergnaud (1985) and Anderson & Ewen (1987).

The plosive represented in (4)b lacks any laryngeal specification. In accordance with a now widespread view, we take this to be the representation of a plain stop (see Harris 1994, Iverson & Salmons 1995, Jessen 1997 and the references there). The status of vocal-fold activity in plain stops varies according to phonological context. Utterance-initially the stops are characterised by zero or short-lag voice onset time; intervocalically they are susceptible to the passive interpolation of vocal-fold vibration between the flanking vowels; utterance-finally they are characterised by a decay in vocal-fold vibration from a preceding vowel (Westbury & Keating 1986). Implicit in the laryngeal non-specification of plain stops is the assumption that the signal correlates of these vocal-fold effects have no informational value. If this is correct, the periodicity of ambient voicing in intervocalic plain stops has the same background status as periodicity in sonorants.

Of course certain laryngeal accompaniments of stops do have informational value and thus must have categorial representation. The locally contrastive properties of aspiration (in English and Danish for example) and active prevoicing (in French and Polish for example) enjoy this status (Harris, in press).

An important implication of these laryngeal considerations for the analysis of lenition is that the process of intervocalic voicing (traditionally known as SONORISATION) cannot be treated on a par with effects such as spirantisation and vocalisation. The latter have a direct impact on the informational content of a segment, since they remove independent cues from the speech signal. Intervocalic voicing, on the other hand, has to be regarded as a secondary consequence of some primary readjustment in the manner categorisation of a segment. In the case of tapping, for example, the loss of a sustained occlusion helps create the aerodynamic conditions for ambient voicing to take place.

4.2 Modelling lenition

A comparison of the specifications of *p* and *w* in b gives a preliminary indication of how the cue-based element model allows us to view the information-degrading impact of lenition on the speech signal as having an analogous impact on phonological representations. With additional place elements, the specific vocalising relation between *p* and *w* generalises to other place series, including the coronal (tapped) and dorsal reflexes illustrated in the Ibibio forms in .

Still more generally, the ability of the model to reflect information-degrading effects extends to all of the major types of lenition (Harris 1990, 1997). With labial place again serving as an exemplar, (5) lists representations of the various lenition reflexes of a plain plosive.

(5)	Plain labial plosive	[edge, noise, rump]
	Labial fricative	[noise, rump]
	Unreleased labial stop	[edge, rump]
	Glottal stop	[edge]
	Glottal fricative	[noise]
	Labial approximant	[rump]

(6) summarises the main types of lenition and their information-degrading effects.

(6)

Lenition type	Example	Suppressed signal cue(s)	Suppressed element(s)
Spirantisation	*p > f*	Silent interval	[edge]
Spirant debuccalisation	*f > h*	Spectral peak	[rump]
Loss of release	*p > p*˺	Release burst	[noise]
Stop debuccalisation	*p*˺ *>* ʔ	Spectral peak	[rump]
Vocalisation	*p > w*	Release burst Silent interval	[noise] [edge]

Each type of lenition suppresses a particular set of frequency- and/or time-domain properties from the speech signal (Lindsey & Harris 1990, Harris & Lindsey 1995). This is matched by the suppression of a particular set of elements. In other words, a reduction in the degree of informational complexity in the speech signal is matched by a reduction in phonological complexity.

As noted earlier, in spite of the procedural flavour of traditional terms such as SPIRANTISATION, DEBUCCALISATION, and VOCALISATION, it is quite possible to analyse synchronic lenition effects in a non-derivational manner (cf. Kirchner 1998, Harris, in press). Phonologically, lenition consists in the exclusion of a given set of elements from a given set of contexts, a state of affairs that can be expressed over output. This can be couched in Optimality-theoretic terms as a competition between two types of output constraint: positional faithfulness constraints requiring the preservation of elements in one context outrank general markedness constraints penalising the appearance of elements in any context (cf. Beckman 1967, Zoll 1998). In a leniting context such as intervocalic position, a vocalised output will be preferred if constraints banning [edge] and [noise] are ranked higher than a constraint banning [rump]. This is tabulated in (7) (* indicates a constraint violation).

(7)

	NO [edge]	NO [noise]	NO [rump]
p	*	*	*
f		*	*
h		*	
p˺	*		*
ʔ	*		
☞ *w*			*

In ‘phonetically driven’ Optimality Theory, markedness constraints delivering lenition effects are claimed to have a functional basis in the principle of least effort (Flemming 1995, Kirchner 1998). Irrespective of what the motivation might be for constraints which suppress segmental categories, it can be readily shown that a model incorporating cue-based elements is better equipped than is orthodox feature theory to express the impact that lenition has on the information-bearing capacity of phonological forms. Compare the element-based account of different lenition types, summarised in (7), with one based on (mostly) standard articulatory features, summarised in (8). (Loss of release (8)c can only be handled through recourse to a non-SPE feature such as [release] or some equivalent (cf. McCawley 1967, Selkirk 1982, Steriade 1993).)

(8)

	Lenition type	Example	*FeatureValue	Output
(a)	Spirantisation	*p > f*	[–continuant]	[+continuant]
(b)	Spirant debuccalisation	*f > h*	Place [+consonantal]	Ø [–consonantal]
(c)	Loss of release	*p > p˺*	[+release]	[–release]
(d)	Stop debuccalisation	*p˺ > ʔ*	Place [+consonantal]	Ø [–consonantal] [+constricted]
(e)	Vocalisation	*p > w*	[–continuant] [+consonantal] [–sonorant] [–voice]	[+continuant] [–consonantal] [+sonorant] [+voice]

The column labelled *FeatureValue in refers to specifications targeted by the family of lenition-delivering constraints. The failure of the standard feature account to capture the informational consequences of vocalisation, set out in e, has already been discussed at length above. The other types of lenition admittedly do not fare quite so badly under this account, partly because fewer feature values are implicated. Deletion of a geometric Place node (indicated by Ø in (8)b and (8)d) certainly indicates a loss of information. However, all other cases involve the switching of the plus-minus value of each feature — in OT terms, the selection of a candidate bearing a given value in preference to a candidate bearing the complement value. This implies the replacement of one piece of information by another, thereby missing the point that a lenited output is informationally impoverished compared to an unlenited congener.

5. The Ibibio foot

5.1 The foot as an informational domain

Information is not evenly distributed across phonological strings, its occurrence being subject to segmental, prosodic, or morphological conditions. Rich informational content is typically concentrated in positions of prosodic or morphological prominence — a fact captured in recent output-oriented theory by means of positional faithfulness constraints of the type referred to above. Informationally impoverished positions, such as those displaying neutralisation, typically occur in contexts that are prosodically weak or morphologically recessive (affixes for example).

One prosodic domain that hosts informational asymmetries is the metrical foot. Across different languages, it is frequently observed that a full vowel inventory is only supported in the head nucleus of a foot, with contracted subsystems showing up in weak nuclei (English being a textbook example.) The imbalance extends to consonantal positions: in one recurrent pattern, the system of consonantal contrasts in a language is maximally displayed in the onset of the head syllable of the foot but is reduced elsewhere (see Harris 1997 for examples and discussion). In this section, we present evidence supporting the conclusion that the foot plays a central role in regulating informational asymmetries in Ibibio.

The most familiar criterion for foot-hood is stress prominence, a property absent from Ibibio, a lexical-tone language. On the other hand, Ibibio does possess at least two other properties that can be considered diagnostic of feet (cf. Hayes 1995). It displays syllable-quantity restrictions and segmental-distributional asymmetries that are strongly reminiscent of foot-based patterns encountered in languages with stress prominence.

5.2 The foot as a stem template

The quantitative restrictions are templatic in nature. They place limits on the size and shape of the Ibibio 'inflectional stem', which consists of a verb root plus an optional suffix. (INFLECTIONAL STEM is the term Hyman (1990) applies to the cognate pattern found in the closely related language Efik.) The basic template takes the form CVXCV, where X stands for either V or C. The resemblance to a heavy-light trochee is unmistakable, as Akinlabi & Urua (1992, 1994) have pointed out. Similar observations have been made with respect to Efik (Cook 1985, Hyman 1990). In fact, foot-templatic verbal morphology seems to be an areal characteristic of the genetically diverse languages spoken in the Cameroon-Nigeria border region (Hyman 1990).[6] Extensive discussion and exemplification of this phenomenon in Ibibio are to be found in Urua (1990), (1999) and Akinlabi & Urua (1992, 1994).

In Ibibio, the CVXCV template places an upper bound on the size of the inflectional stem. (9) illustrates the six attested canonical stem shapes that are contained within this limit.

(9)

(a)	[CV]	*dí*	*sé*	*dá*
		'come'	'look'	'stand'
(b)	[CVC]	*kàt*	*tɔ́k*	*tóp*
		'show'	'urinate'	'throw'
(c)	[CVCV]	*kárá*	*séɣé*	*bɔ́ɣɔ́*
		'govern'	'be childish'	'overtake'
(d)	[CVVC]	*dííp*	*kɔ́ɔ́t*	*bùùk*
		'hide'	'call'	'bury'
(e)	[CVCCV]	*bèkké*	*bɔ̀kkɔ́*	*bʌ̀kkɔ́*
		'belch'	'escape'	'uproot'
(f)	[CVVCV]	*sèèmé*	*tòòró*	*sɔ́ɔ́ŋɔ́*
		'bemoan'	'praise'	'affirm'

The template also acts as a lower bound for certain verbal paradigms. In this case, potentially oversized or undersized morphological material is tailored to a fixed

[6] For detailed treatments of foot-templatic effects in other languages in this area, see Hyman on Gokana (1982, 1985) and Basaa (2000).

CVXCV template through segment truncation or augmentation.[7] These effects are illustrated by the negative, frequentative and reversive forms in (10), (11) and (12).

(10) NEGATIVE: root + *ké*

(a)	CVC				
		tèm	'cook'	*í-tèmmé*	'is not cooking'
		kɔ̀p	'lock'	*í-kɔ̀ppɔ́*	'is not locking'
		fʌ́k	'cover'	*í-fʌ́kkɔ́*	'is not covering'
(b)	CV				
		sé	'look'	*í-séé*ɣ*é*	'is not looking'
		nɔ̀	'give'	*í-nɔ̀ɔ̀*ɣ*ɔ́*	'is not giving'
		dá	'stand'	*í-dáá*ɣ*á*	'is not standing'
(c)	CVVC				
		síít	'block'	*í-síí*ɾ*é*	'is not blocking'
		fáák	'wedge'	*í-fáá*ɣ*á*	'is not wedging'
		kɔ́ɔ́ŋ	'hang on hook'	*í-kɔ́ɔ́ŋɔ́*	'is not hooking'

(11) FREQUENTATIVE: root + N*é*

CV			
nɔ̀	'give'	*nɔ̀ŋŋɔ̀*	'give (freq.)'
kpá	'die'	*k͡páŋŋá*	'die (freq.)'

(12) REVERSIVE: root + C*é*

CVVC			
síít	'block'	*sítté*	'unblock'
fáák	'wedge'	*fákká*	'remove wedge'
kɔ́ɔ́ŋ	'hang on hook'	*kɔ́ŋŋɔ́*	'unhook'

Stems consisting of a CVC root and a CV suffix satisfy the fixed-templatic restriction by default (see (10)a). Attachment of a CV suffix to CV roots is accompanied by either vowel lengthening (see (10)b) or consonant gemination (see (11)). Suffixation to CVVC roots results either in consonant truncation (see (11)c) or vowel shortening (see (12)).

5.3 Segmental asymmetries within the foot

As to the segmental characteristics of the CVXCV template, it hosts contrastive asymmetries similar to those associated with trochees in languages with stress prominence. In order to describe these effects, it will be useful to make a terminological distinction between the foot HEAD, consisting of the initial CV of the template, and the TAIL, consisting of any residual positions. Only the head sponsors the full set of vowel and consonant distinctions in Ibibio. The contrastive potential of the tail is greatly curtailed: not only does it lack a proportion of the segmental material available to the head, but what material it does have is to a great extent assimilated from the head.

[7] As with SPIRANTISATION and VOCALISATION discussed above, the use of terms such as TRUNCATION and AUGMENTATION is not intended to imply a derivational analysis. In keeping with a fully output-oriented approach, they can be understood here in a purely descriptive sense to refer to cross-paradigm comparisons.

In the case of syllable nuclei, assimilatory neutralisation in the foot tail manifests itself in two ways. First, as indicated by the examples in b, c and , V_2 in a V_1V_2 cluster is invariably a copy of V_1, both tonally and segmentally. Second, a stem-final vowel occurring in the tail is harmonically dependent on the head nucleus. This effect, already suggested by some of the examples above, is more fully exemplified by the frequentative and relative forms in (13). Here we see how the quality of the tail-final vowel is determined by the head nucleus in terms of frontness, roundness and ATR.

(13)	(a)	*díí-mé*	'lift up'	*kúú-mɔ́*	'open'
		fèè-ŋé	'run'	*dóó-mó*	'become heavy'
		kpáŋ-ŋá	'die'	*dɔ́ɔ́-mɔ́*	'wipe up'
	(b)	*áà-sììɾè*	'who blocks'	*áà-dùùhò*	'who is alive'
		áà-sèèhè	'who looks'	*áà-dòmmò*	'who bites'
		áà-dààhà	'who stands'	*áà-kɔ̀ppɔ̀*	'who locks'

Non-nuclear positions within the tail are also subject to neutralisation, failing to sustain a proportion of the laryngeal and manner contrasts borne by onsets within the head. The contrastive asymmetries between head and tail consonants are set out in (14), which summarises the distribution of oral stops and related segments in Ibibio.

(14)

Foot head		Foot tail		
[C		VCCV	VC] {‖/C}	VC(])V
k͡p	*b*	*pp*	*p*	*ß*
t	*d*	*tt*	*t*	*ɾ*
k		*kk*	*k*	*ɣ*

The laryngeal contrast holding in head onsets takes the form of a distinction between plain (voiceless unaspirated) and prevoiced plosives (see Connell 1991, Urua 1999, Harris, in press). In tail-internal geminates, this contrast is suspended in favour of plain stops.

Singleton consonants within the tail are subject to the lenition effects already sampled in and now more fully illustrated in (15).

(15)	(a)	*dɨ́p*	'hide'	*dɨ́ßé*	'hide oneself'
		déép	'scratch'	*dééßé*	'not scratching'
		bɔ́p	'tie'	*bɔ́ßɔ́*	'tie oneself'
		tóp	'tie'	*tóßó*	'tie oneself'
		bèt	'shut'	*bèɾé*	'be shut'
		kɔ́ɔ́t	'call'	*kɔ́ɔ́ɾó*	'not calling'
		sɨ́ɨ́t	'block an opening'	*sɨ́ɾé*	'be blocked'
		fʌ́k	'cover'	*fʌ́ɣɔ́*	'cover oneself'
		fááḱ	'wedge'	*fááɣá*	'not wedged'
		bɔ́k	'group together'	*bɔ́ɣɔ́*	'be grouped together'
	(b)	*kɔ̀p*	'lock'	*kɔ̀ß úsʌ́ŋ*	'lock the door'
		bèt	'push'	*bèɾ ɔ́wɔ́*	'push someone'
		kʌ̀k	'shut'	*kʌ̀ɣ úsʌ́ŋ*	'shut the door'

The alternating pairs illustrated in (15) are the only oral consonants permitted tail-finally in Ibibio. (Nasal stops are also allowed in this position.) Utterance-finally or before a word-initial consonant, the segments are realised as stops. Impressionistically described, they tend to be unreleased and characterised by rapid decrescendo voicing from the preceding vowel. Intervocalically, they lenite in the manner detailed above. As a comparison of (15)a and (15)b shows, the triggering vowel may or may not be separated from the target consonant by a word boundary.[8]

To lenite, an Ibibio consonant must appear in the tail of a foot. The necessity of this condition is confirmed by the fact that the presence of a following vowel is not in itself a sufficient condition for lenition to take place. An intervocalic consonant resists lenition if it either occupies a foot head or falls outside a foot. In the examples in (16), the prevocalic context is provided by a prefix vowel (nominalising in (16)a, pronominal in (16)b). The consonants following the prefix are initial in the root and thus also in the foot. Occupancy of a head position grants these intervocalic consonants immunity to lenition.

---

[8] In the absence of an intervocalic contrast between geminate and non-geminate plosives here, it might initially seem reasonable to treat the plosives as singleton consonants (cf. Connell 1991). However, this would disturb the otherwise uniform quantitative patterning of verbal paradigms where we have independent evidence that the heavy-light trochee defines a fixed prosodic template. In the case of the negative paradigm, for example, a CVCV analysis of forms containing an intervocalic plosive would sever the quantitative link with clear cases of CVXCV involving either a long vowel (as in *kɔ́ɔ́ɾó* 'not call') or a geminate nasal (as in *dɔ́mmɔ́* 'not bite'). Geminate nasals contrast with singletons intervocalically, e.g. *dɔ́mmɔ́* 'not bite' versus *dɔ́mɔ́* 'switch on light'.

On reasons for rejecting the view that intervocalic leniting consonants are ambisyllabic in Ibibio, or indeed in any language, see Harris (in press).

(16)	(a)	*ú*-[*táŋ*]	**úɾáŋ*	'plaiting'
		ú-[*kʌ́p*]	**úɣʌ́p*	'covering'
	(b)	*í*-[*bàttá*]	**íßàttá*	'(s)he is not counting'
		á-[*tòòró*]	**íɾòòró*	'(s)he is praising'

To see what happens to consonants that are intervocalic but extra-pedal, consider the examples containing the negative/reversive suffix *ké* in (17). In (17)a, the suffix is enclosed within the template of an inflectional stem. The suffix consonant thus occupies a foot tail and undergoes lenition in the expected manner.

(17)	(a)	[*séé-ɣé*]		'not look'
		[*dáá-ɣá*]		'not stand'
	(b)	[*dáppá*]-*ké*	**dáppáɣá*	'not dream'
		[*kɔ́ŋŋɔ́*]-*ké*	**kɔ́ŋŋɔ́ɣɔ́*	'not unhook'
		[*dámmá*]-*ké*	**dámmáɣá*	'is not crazy'
		[*fááŋá*]-*ké*	**fááŋáɣá*	'not argue'

In the examples in (17)b, on the other hand, the same suffix is attached to a verbal template that is already saturated by maximal CVCCV or CVVCV material. Excluded from the foot, the suffix consonant is not subject to lenition. Note that vowel harmony too is foot-sensitive: as the examples in b demonstrate, a foot-external suffix vowel does not harmonise with a root vowel.

Strong support for the existence of feet in Ibibio comes from the observation that the domain-sensitivity of lenition illustrated in (17) is, aside from stress prominence, in all significant respects identical to what can be found in languages with stress feet. Focusing on coronals, compare the conditions on prevocalic tapping in Ibibio with those in English:

(18)			English	Ibibio	
	(a)	Stop	bòu[tíque]	*ú*[*táŋ*]	'plaiting'
	(b)	Tap	[gèt] óff [gét] a	[*bèɾ*] *ówó*	'push someone'

In English, stress is only tangentially implicated in tapping (despite what is often assumed; cf. Ladefoged 2001: 58). As the English examples in (18)b demonstrate, word-final *t* taps regardless of whether the following vowel bears stress or not. What is crucial is the consonant's location with respect to foot structure: just as in Ibibio, tapping occurs foot-finally (see (18)b) but not foot-initially (see (18)a). (For foot-based analyses of English tapping, see Kiparksy 1979, Harris 1997 and Jensen 2000.)

5.4 Distribution of elements within the foot

In summary, there is strong evidence that Ibibio morphology is built around a foot-sized template. Within this domain, there exists a clear segmental-informational imbalance: distinctions sustained in the head CV of the foot are neutralised in the foot tail.

As an initial hypothesis, the manifestation of this informational asymmetry in singleton consonants can be characterised in terms of the cue-based elements displayed in (19).

(19)

		Foot head C	Tail C
(a)	[prevoice]	✓	✗
(b)	[edge]	✓	✗ (_V)
(c)	[noise]	✓	✗

Active prevoicing in obstruents is permitted in the onset of the foot head but not elsewhere ((19)a). An interval of radically reduced amplitude is supported in the foot head but not in the foot tail if a vowel follows ((19)b). The foot head can host aperiodic energy, reflecting either continuous frication or plosive release ((19)c). There are two respects in which aperiodic energy might be said to be missing from the foot tail — firstly in that fricatives are barred from this site, and secondly to the extent that it is accurate to describe lenited consonants as frictionless continuants when prevocalic and as unreleased stops elsewhere in this context.

The characterisation of the informational asymmetries in is based on a combination of impressionistic observation and preliminary acoustic analysis (Connell 1991, Urua 1999, Harris, in press). In the next section, we attempt to quantify the extent to which the asymmetries are reliably detectable in the speech signal.

6. Lenition and the speech signal

6.1 Extracting edges and noise from speech signals

The quantitative study reported in this section has two main objectives. One is primarily descriptive — to provide a more accurate specification of the reportedly continuous effects of Ibibio lenition than is possible with impressionistic phonetic labelling. The other is to bring experimental evidence to bear on the posited relation between properties of the speech signal and the cue-based elements described in §3.

More specifically, the aim of the study is to address the following interlocking questions. How much of the signal information projected by unlenited consonants in Ibibio is retained by their lenited counterparts? In particular, to what extent do consonants in leniting contexts display the reduced-amplitude interval of stops and the aperiodic energy of plosion or continuous frication? How good is the fit between the detection of these particular properties in the speech signal and the hypothesised occurrence of their element correlates, [edge] and [noise]? To what extent can different positions within the Ibibio foot be reliably distinguished on the basis of these properties?

The data are drawn from audio recordings of four adult native speakers of Ibibio (two female, two male). The subjects produced a word set containing all of the stops and related reflexes shown in , located within a representative sample of phonological contexts presenting different foot positions and different following vowels. The present study is based on a sample of 400-plus word tokens, roughly equally distributed across the four speakers.

For each word token selected for analysis, energy and aperiodicity measurements were taken within a frame containing a target consonant preceded by a vowel and followed by a sonorant (vowel or nasal consonant).[9] In the case of word-final consonants, the sonorant was supplied by a following word in a carrier phrase.[10] Measurement commenced at the mid point of the pre-target vowel and ended at the mid point of the post-target sonorant. The following foot positions, differentiated on the basis of the morphological paradigm criteria discussed in §5, are represented in the data: foot-initial (V[CV), foot-medial singleton consonants (VCV) and geminates (VCCV), foot-final prevocalic (VC]V), and foot-final preconsonantal (VC]C). The full set of contexts represented by the VCX frame is set out in (20), together with examples and a reference key to be used in the diagrams below.

(20)

	Key	Foot context	Example	
(a)	Ai	V[CV	*ìtà*	'punch'
(b)	Am	VCCV	*sákká*	'be split'
(c)	Af	VC]C	*tóp ǹnɔ̀*	'..throw for me'
(d)	Bm	VCV	*táβá*	'forfeit'
(e)	Bf	VC]V	*sáɾ á*	'discriminate'

The reference key in reflects the distinction, drawn on the basis of the impressionistic observations in §5, between foot sites resistant to spirantisation/vocalisation (the 'A contexts' Ai, Am, Af) and those prone to it (the 'B contexts' Bm, Bf).

The analyses to be presented below investigated the extent to which the data can be classified on the basis of [edge] (§6.2) and [noise] (§6.3).

6.2 Edges

As described in §4, the hypothesised acoustic signature of the [edge] element is an abrupt and sustained reduction in amplitude. To determine the extent to which the data can be classified on the basis of this property, an algorithm was devised which measures the fluctuation of acoustic energy across a given analysis frame.

Within the VCX analysis frame, energy was sampled every 5ms within a 30ms-sized window in each of four frequency bands: 100Hz-5kHz (overall amplitude), 100Hz-2kHz (low), 1.5-3.5kHz (mid) and 3-5kHz (high).[11] Only the results for the overall and low bands are reported below, since these were found to deliver the most reliable classification of the data. For each word token, standard deviation values were

---

[9] Analysis of the speech data was performed using the SFS software designed by Mark Huckvale at University College London (http://www.phon.ucl.ac.uk/resource/sfs/).

[10] The carrier phrase is *ḿbô _ ǹnɔ̀* 'I say _ for myself'.

[11] In order to exclude low-frequency rumble, no energy measurements were taken below 100Hz.

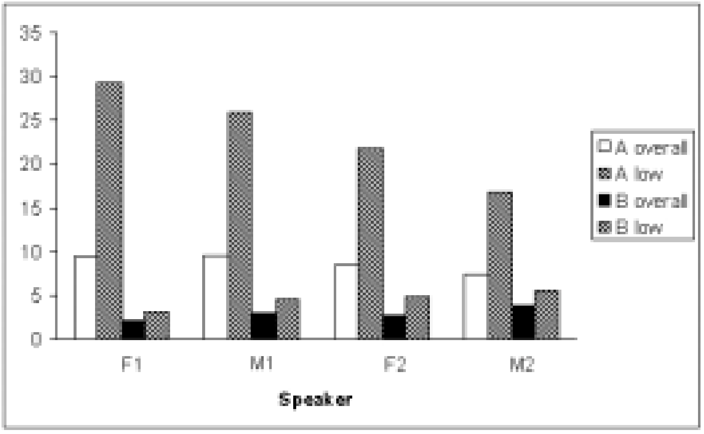
computed (in dB) for each frequency band; the higher the value, the greater the degree to which energy varies across the time interval of an analysis frame. Since the segments on both sides of the target site are high-energy sonorants, a relatively high standard deviation value can reasonably be taken to reflect a drop in energy during the target consonant.[12]

The results, summarised in Figure 1, reveal systematic differences between the A and B contexts. For all speakers, the mean values of the two contexts are different at a high level of statistical significance in both the overall-frequency and the low-frequency dimensions.[13] Moreover, the general directionality of the differences bears out the expectation that the lenition-favouring B contexts should show lower energy values, reflecting a relatively shallow dip in amplitude during the transition between the target consonant and the flanking sonorants.

[12]No tokens showed evidence of an energy INCREASE during the target segment.

[13]The significance levels, in every case $p < .0001$, were calculated using a one-tail t-test (not assuming equal values), with the following results:

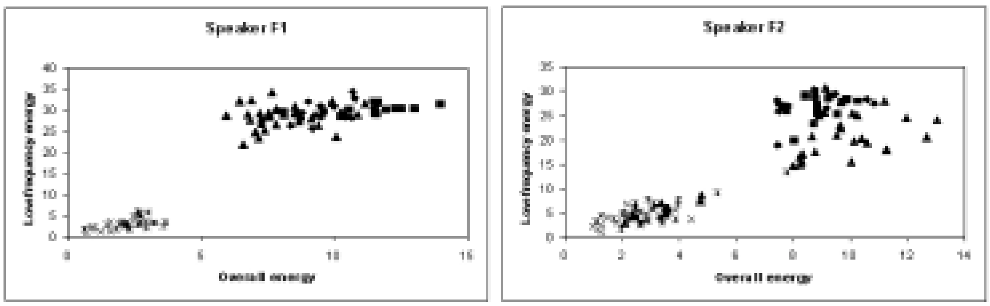
Speaker		Overall energy		Low-frequency energy	
		A	B	A	B
F1	Mean	9.29	2.21	29.22	3.24
	t stat	25.65		68.38	
M1	Mean	9.55	2.91	25.85	4.49
	t stat	18.09		50.89	
F2	Mean	8.44	2.76	21.77	4.86
	t stat	16.41		16.27	
M2	Mean	7.36	3.77	16.80	5.51
	t stat	7.63		15.27	



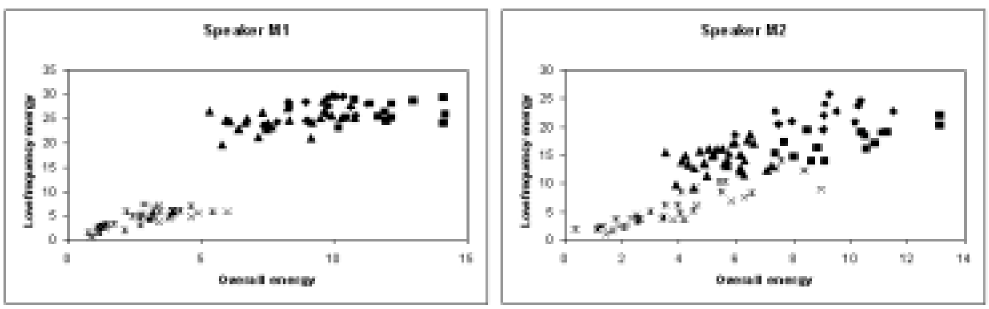
N		F1	M1	F2	M2
Context A	V[CV, VCCV, VC]C	62	52	63	61
Context B	VCV, VC]V	35	43	47	38

Figure 1. *Ibibio consonants: mean energy values (standard deviations, dB) for two frequency bands (overall and low) in two sets of phonological contexts.*

The dispersion of values is displayed in Figure 2, where the data is cross-classified by overall energy and low-frequency energy. Here we see distributions that are consistent with earlier reports of inter- and intra-speaker variability in Ibibio lenition. The clearest picture is represented in Figure 2a and 1c, in which the data for speakers F1 and M1 divide neatly along the lines of the A-versus-B contextual distinction. The separation is particularly robust in the case of Figure 2a, where either frequency dimension on its own would be sufficient to provide a discrete classification of the data.



(a) Speaker F1 (b) Speaker F2

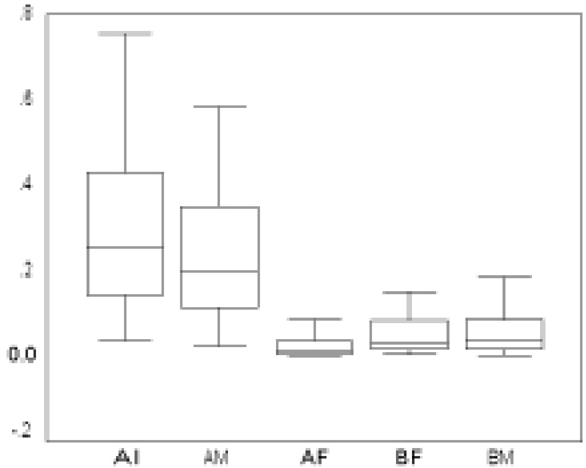


(c) Speaker M1 (d) Speaker M2

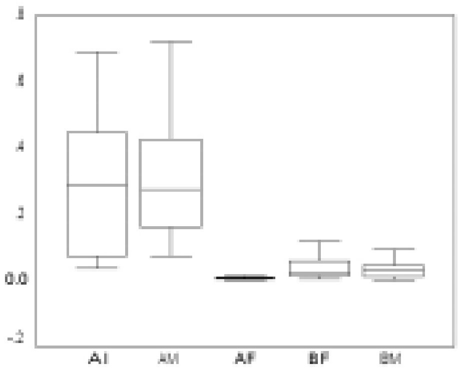
Contexts						
	◆	Ai	V[CV			
	■	Am	VCCV	×	Bm	VCV
	▲	Af	VC]C	✳	Bf	VC]V

Figure 2*. Energy fluctuation across Ibibio vowel-consonant-sonorant sequences in five phonological contexts: standard deviation values (in dB) for overall-frequency (100Hz-5kHz) and low-frequency bands (100-Hz-2kHz).*

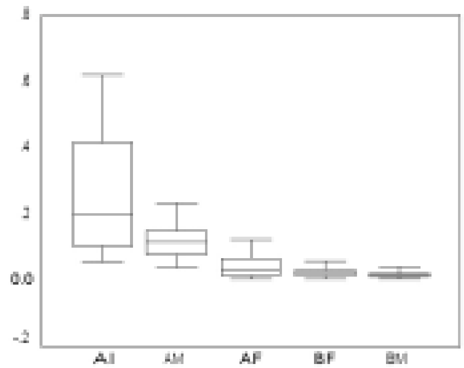
Things are less clear-cut in Figures 2b and 2c, where the A and B data can be seen to overlap, though in different ways. The pattern presented by speaker F2 (Figure 2b) is like those of F1 and M1 to the extent that the data bifurcate neatly. However, in this case, there is no perfect correlation with the A–B contextual distinction, a handful of rogue tokens showing up in the 'wrong' region. Most of these (nine out of 34 Af tokens) involve foot-final preconsonantal segments with unexpectedly low values, indicating lenition in a context that otherwise resists it. In Figure 2d, in contrast, the data for speaker M2 is arrayed along a continuum; while the means of the A and B tokens are clearly distinct, there is a region where the two sets of data overlap. These different patterns of overlap conform to two of the major types of phonological variability reported in the language-variation literature, one involving phonetic gradience, the other involving alternation between phonetically discrete variants (see Labov 1994 for discussion and references).



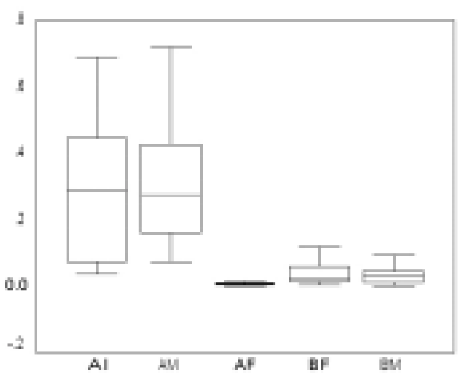
(a) All speakers



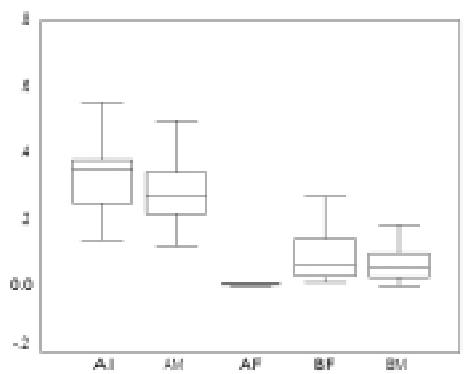
(b) Speaker F1



(c) Speaker F2



(d) Speaker M1



(e) Speaker M2

Contexts	Ai	V[CV		
	Am	VCCV	Bm	VCV
	Af	VC]C	Bf	VC]V

***Figure 3**. Normalised noise values (aperiodicity * energy) for Ibibio consonants in five phonological contexts (median, interquartile range, high and low values).*

6.3 Noise

Aperiodic energy is the acoustic signature hypothesised for the element [noise] (see §4). The algorithm designed to gauge the classifiability of the Ibibio data on the basis of this property provides a means of relating separate measurements for energy and aperiodicity.

The analysis calculates the degree of aperiodicity in a signal as a function of time by using an autocorrelation measure to estimate what proportion of the signal is predictable at each time instant.[14] The output value, normalised to overall energy, is then related to a normalised energy value (both expressed as simple fractions).[15] The algorithm finds the time in the analysis frame with the minimum energy, in order to locate the target consonant, and then searches forward until it locates the maximum product of aperiodicity and energy. The resulting maximum product values provide the data for Figure 3: the higher the value, the greater the degree of aperiodic energy in the signal.

The results, diagrammed in Figure 3, indicate that foot context has a highly significant effect on the overall distribution of noise values for all speakers ($p < .0001$, ANOVA single factor). As with the edge analysis, there is some degree of overlap between individual contexts. Nevertheless, there is an obvious distinction between Ai/Am with relatively high values on the one hand and Af/Bf/Bm with low values on the other. The difference between the means of these two groups is highly significant ($p <.0001$, pairwise two-tailed t-tests).

The noise results for the B data are revealing for the light they shed on the nature of so-called spirantisation. The low values displayed by Bf and Bm on the edge analysis confirm these as lenition-favouring contexts. The fact that they also display low values on the noise analysis suggests that the lenition does not produce canonical fricatives, indicating that VOCALISATION is a more appropriate term here than SPIRANTISATION.

Also noteworthy is the behaviour of Af, the foot-final preconsonantal context. Under the edge analysis, this context patterns with the other A data in showing a marked drop in energy, indicative of a maintenance of stop closure during the consonant. Under the noise analysis, on the other hand, Af patterns with the B data in showing significantly lower values than the other A data. This confirms the tendency for final stops not to be characterised by a noisy release burst.

6.4 Discussion

The edge and the noise analyses provide intersecting two-way classifications of the data by phonological context, the separation being more clear-cut for certain speakers and contexts than others.

The measurement of energy fluctuation across the VCX analysis frame divides the contexts along the following lines. Foot-initial (Ai), geminate (Am), and preconsonantal (Af) consonants show relatively high levels of fluctuation, indicating

---

[14]The measure is probably irrelevant in silent intervals.

[15]The normalisation procedure expresses the energy of all points within the analysis frame as a simple fraction of the point with the largest energy.

resistance to lenition. The degree of energy fluctuation in prevocalic singleton consonants in the foot tail (Bm and Bf) is significantly lower, confirming these as lenition targets.

Measuring the degree of aperiodic energy in the VCX frame yields a different two-way classification. Geminates (Am) and singleton consonants in the foot head (Ai) show relatively high levels of aperiodic energy. Since the edge analysis establishes these contexts as resistant to lenition, the energy can reasonably be attributed to noise accompanying plosive release. Singleton consonants in the foot tail, in contrast, show significantly lower levels of aperiodic energy, indicating a relative lack of continuous frication noise in lenited segments (Bm, Bf) and of burst release in unlenited segments (Af).

The results confirm that, taken in the round, the effects of stop allophony in Ibibio are variable and phonetically continuous. Nevertheless, they also indicate that they reach near-categoriality for some speakers in certain phonological contexts. The analyses deliver a reasonable fit between the signal properties they target and the hypothesised occurrence of the phonological elements [edge] and [noise]. In particular, they confirm the hypothesis that information is asymmetrically distributed across the foot. We now have instrumental evidence that the informational asymmetries sketched in (19) do indeed leave their mark on the speech signal. Specifically, singleton consonants in the foot tail preferentially lack at least two pieces of information freely supported by consonants in the foot head — the aperiodic energy associated with the [noise] element and, when prevocalic, the radical amplitude drop associated with the [edge] element.

It is for future perception research to determine whether these foot-based informational asymmetries can be exploited by Ibibio listeners. In particular serve as parsing cues to prosodic and thus morphological domain structure.

7. Conclusion

SPE-style feature specification portrays a phonological string as a steady stream of information in which each segment bears as much information as the next. A representation consisting of cue-based elements presents a quite different picture, an uneven one in which the information flow is heavy in certain string positions and light in others. This scenario gives direct expression to the observation that contrastive potential, an important — albeit not exclusive — diagnostic of informational yield, is unequally distributed across different phonological contexts. Moreover, compared to SPE feature theory, it defines a more transparent relation between phonology and informational asymmetries in the speech signal.

Contextually determined consonantal lenition is symptomatic of a more general effect whereby segmental information is attracted to positions of prosodic prominence. Indeed there is a case for saying that segmental content contributes as much to the relative prominence of a position as properties more traditionally associated with that function, such as stress accent or pitch accent (cf. Hayes 1995). If this is correct, then a tone language such as Ibibio shows how segmental asymmetries can signal prominence relations even in the absence of stress.

Finally, returning to a theme touched on at the beginning of this paper, note that the model presented here is neutral on the issue of whether articulatory effort exerts an active influence on phonologically entrenched lenition. We have good evidence that articulatory gestures are less extreme in prosodically weak positions (see Pierrehumbert & Talkin 1992 and de Jong 1998 for discussion of the relevant literature and exemplification). However, arguing that lenition in weak contexts results from a speaker-oriented pressure to minimise the expenditure of articulatory effort is teleologically problematic. The effect could equally well reflect a listener-oriented pressure to provide syntagmatically useful cues to the asymmetric distribution of phonological information.

References

Abramson, A. S. & A. Lisker (1970). Discriminability along the voicing continuum: cross-language tests. *Proceedings of the 6th International Congress of Phonetic Sciences*, 569-573. Prague: Academia.

Akinlabi, A. & E. E. Urua (1992). Prosodic target and vocalic specification in the Ibibio verb. In J. Mead (ed.) *WCCFL11: Proceedings of the 11th West Coast Conference on Formal Linguistics*. Stanford, CA: CSLI. 1-14.

Akinlabi, A. & E. E. Urua (1994). Foot in the Ibibio verb. Ms, Rutgers University and University of Uyo.

Anderson, J. M. & C. Jones (1974). Three theses concerning phonological representations. *Journal of Linguistics* 10. 1-26.

Anderson, J. M. & C. J. Ewen (1987). *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Archangeli, D. & D. Pulleyblank (1994). *Grounded phonology*. Cambridge, MA: MIT Press.

Atal, B. S., J. J. Chang, M. V. Matthews & J. W. Tukey (1978). Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer sorting technique. *Journal of the Acoustical Society of America* 47, 65 (A).

Beckman, J. N. (1997). Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology* 14. 1-46.

Blumstein, S. E. & K. N. Stevens (1979). Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America* 66. 1001-1017.

Blumstein, S. E. & K. N. Stevens (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America* 67. 648-662.

Blumstein, S. E. & K. N. Stevens (1981). Phonetic features and acoustic invariance in speech. *Cognition* 10. 25-32.

Boersma, P. (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. LOT dissertations 11. The Hague: Holland Academic Graphics.

Borowsky, T., J. Itô & R-A. Mester (1984). The formal representation of ambisyllabicity: evidence from Danish. *North Eastern Linguistics Society* 14. 34-48.

Brockhaus, W. & M. Ingleby (in press). Phonological primes: cues and acoustic signatures. In J. Durand & B. Laks (eds.), *Phonology: from phonetics to cognition*. Oxford: Oxford University Press.

Brockhaus, W., M. Ingleby & C. Chalfont (1994). Robust techniques for recognition of new knowledge-based speech primitives. In R. Togneri (ed.), *Proceedings of the 5th Australian International Conference on Speech Science and Technology*, vol 1, 374-380. Canberra: Australian Speech Science and Technology Association.

Browman, C. P. & L. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* 6. 201-52.

Burton-Roberts, N, P. Carr & G. Docherty (eds.) (2000). *Phonological knowledge: its nature and status*. Oxford: Oxford University Press.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2. 223-250.

Clements, G. N. & Susan R. Hertz (1991). Nonlinear phonology and acoustic interpretation. *Proceedings of the XIIth International Congress of Phonetic Sciences*, Vol 1/5, 364-373. Provence: Université de Provence.

Clements, G. N. & E. V. Hume (1995). The internal organization of speech sounds. In J. A. Goldsmith (ed.). *The handbook of phonological theory*, 245-306. Oxford: Blackwell.

Cole, J., J. I. Hualde & K. Iskarous (1998) Effects of prosodic and segmental context on /g/-lenition in Spanish. In O. Fujimura, B. Joseph & B. Palek (eds.), *Proceedings of LP98*, 575-589. Prague: Karolinium.

Connell, B. (1991). Phonetic aspects of the Lower Cross languages and their implications for sound change. PhD dissertation, University of Edinburgh.

Cook, T. L. (1985). An integrated phonology of Efik. Vol 1. PhD dissertation, Universiteit Leiden.

Davidsen-Nielsen, N. (1978). *Neutralization and archiphoneme: two phonological concepts and their history*. Copenhagen: Akademisk Forlag.

Dorman, M., M. Studdert-Kennedy & L. Raphael (1977). The invariance problem in initial voiced stops: release bursts and formant transitions as functionally equivalent context-dependent cues. *Perception and Psychophysics* 22. 109-122.

Flemming, E. (1995). Auditory representations in phonology. PhD dissertation, UCLA.

Flemming, E. (in press). Phonetic detail in phonology: towards a unified account of assimilation and coarticulation. To appear in *Phonology* 17.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics* 14. 3-28.

Giegerich, H. J. (1992). *English phonology: an introduction*. Cambridge: Cambridge University Press.

Gussenhoven, C. (1986). English plosive allophones and ambisyllabicity. *Gramma* 10. 119-41.

Hale, M. & C. Reiss (2000). Phonology as cognition. In Burton-Roberts, Carr & Docherty 2000, 161-184.

Halle, M. (1959). *The sound pattern of Russian*. The Hague: Mouton.

Halle, M. & K. N. Stevens (1971). A note on laryngeal features. *Quarterly Progress Report of the Research Laboratory of Electronics* (MIT) 101. 198-213.

Harris, James (1969). *Spanish phonology*. Cambridge, MA: MIT Press.

Harris, James (1984). La espirantisación en castellano y la representación fonológica autosegmental. *Estudis Grammaticals: Working Papers in Linguistics* 1. 149-167. Universitat Autònoma de Barcelona.

Harris, John (1990). Segmental complexity and phonological government. *Phonology* 7. 255-300.

Harris, John (1994). *English sound structure*. Oxford: Oxford University Press.

Harris, John (1997). Licensing Inheritance: an integrated theory of neutralisation. *Phonology* 14. 315-370.

Harris, John (in press). Release the captive coda: the foot as a domain of phonetic interpretation. To appear in *Papers in Laboratory Phonology* 6. Oxford: Oxford University Press.

Harris, John & G. Lindsey (1995). The elements of phonological representation. In J. Durand & F. Katamba (eds.), *Frontiers of phonology: atoms, structures, derivations*. Harlow, Essex: Longman. 34-79.

Harris, John & G. Lindsey (2000). Vowel patterns in mind and sound. In Burton-Roberts, Carr & Docherty 2000, 185-205.

Harris, John & E. E. Urua (1998). Direct phonetic interpretation: foot-based lenition in Ibibio. Paper given at *Current Trends in Phonology II*, Royaumont, 22-24 June 1998.

Hayes, B. (1986). Inalterability in CV phonology. *Language* 62. 321-51.

Hayes, B. (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.

Hyman, L. M. (1982). The representation of nasality in Gokana. In H. van der Hulst & N. Smith (eds), *The structure of phonological representations*. Part I, 111-130. Dordrecht: Foris.

Hyman, L. M. (1985). *A theory of phonological weight*. Dordrecht: Foris.

Hyman, L. M. (1990). Non-exhaustive syllabification: evidence from Nigeria and Cameroon. *CLS* 26: 2. 175-195.

Hyman, L. M. (2000). Basaa. Ms, Berkeley.

Hyman, L. (in press). On the limits of phonetic determinism in phonology: *NC revisited. In E. V. Hume & K. Johnson (eds.), *The role of perception in phonology*. Academic Press.

Iverson, G. K. & J. C. Salmons (1995). Aspiration and laryngeal representation in Germanic. *Phonology* 12. 369-396.

Jakobson, R. (1968). *Child language, aphasia and phonological universals*. (Translated by A. Keiler.) The Hague: Mouton.

Jakobson, R., G. Fant & M. Halle (1952). *Preliminaries to speech analysis*. Cambridge, MA: MIT Press.

Jensen, J. T. (2000). Against ambisyllabicity. *Phonology* 17. 187-235.

Jessen, M. (1997). *Tense vs. lax obstruents in German. Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung* 3. Universität Stuttgart.

Jong, K. de (1998). Stress-related variation in the articulation of coda alveolar stops: flapping revisited. *Journal of Phonetics* 26. 283-310.

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. PhD dissertation, MIT. Published 1980, New York: Garland.

Kaye, J., J. Lowenstamm & J-R. Vergnaud (1985). The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook* 2. 305-328.

Keyser, S. J. & K. N. Stevens (1994). Feature geometry and the vocal tract. *Phonology* 11. 207-236.

Kiparsky, P. (1979). Metrical structure is cyclic. *Linguistic Inquiry* 8. 421-42.

Kirchner, R. M. (1998). An effort-based approach to consonant lenition. PhD dissertation, UCLA.

Kirchner, R. M. (forthcoming). Phonological contrast and articulatory effort. In L. Lombardi (ed.), *Segmental processes in Optimality Theory*. Cambridge: Cambridge University Press.

Klatt, D. (1989). Review of selected models of speech perception. In W. D Marslen-Wilson (ed.), *Lexical representation and process*, 169-226. Cambridge, MA: MIT Press.

Labov, W. (1994). *Principles of linguistic change: internal factors*. Oxford: Blackwell.

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. Chicago, IL: University of Chicago Press.

Ladefoged, P. (2001). *A course in phonetics*. 4th ed. Fort Worth, TX: Harcourt.

Ladefoged, P., J. DeClerk, M. Lindau & G. Papcun (1972). An auditory motor theory of speech production. *UCLA Phonetics Laboratory Working Papers in Phonetics* 22. 48-76.

Ladefoged, P. & I. Maddieson (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lahiri, A. & W. Marslen-Wilson (1991). The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition* 38. 245-294.

Lavoie, L. A. (2000). Phonological patterns and phonetic manifestations of consonant weakening. PhD dissertation, Cornell University.

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

Lieberman, P. & S. E. Blumstein (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.

Lindblom, B., J. Lubker & T. Gay (1979). Formant frequencies of some fixed-mandible vowels and a model of speech-motor programming by predictive simulation. *Journal of Phonetics* 7. 147-162.

Lindsey, G. & J. Harris (1990). Phonetic interpretation in generative grammar. *UCL Working Papers in Linguistics* 2. 355-69.

McCarthy, J. J. (1988). Feature geometry and dependency: a review. *Phonetica* 45. 84-108.

McCarthy, J. J. & A. S. Prince (1986). Prosodic morphology. Reprinted 1999 in J. A. Goldsmith (ed.), *Phonological theory: the essential readings*, 238-288. Oxford: Blackwell.

McCarthy, J. J. & A. S. Prince (1995). Faithfulness and reduplicative identity. In J. N. Beckman, L. W. Dickey & S. Urbanczyk (eds.), *Papers in Optimality Theory*. Amherst: GLSA. 249-384.

McCawley, J. D. (1967). Le rôle d'un système de traits phonologiques dans une théorie de langage. *Langages* 8. 112-123.

Mascaró, J. (1984). Continuant spreading in Basque, Catalan and Spanish. In M. Aronoff & R. T. Oehrle (eds.), *Language sound structure*, 287-298. Cambridge, MA: MIT Press.

Moore, B. C. J. (1989). *An introduction to the psychology of hearing*. 3rd ed. London: Academic Press.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (ed.), *The production of speech*, 189-216. New York: Springer.

Ohala, J. J. (1986). Against the direct-realist view of speech perception. *Journal of Phonetics* 14. 75-82.

Ohala, J. J. (1992). Bibliography. *Language and Speech* 35. 5-13.

Ohala, J. J. (1995). Phonetic explanations for sound patterns: implications for grammars of competence. K. Elenius & P. Branderud (eds.), Proceedings of the 13th International Congress of Phonetic Sciences, vol 2, 52-59. Stockholm: KTH & Stockholm University.

Ohala, J. J. & H. Kawasaki-Fukumori (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. the shape of morphemes. In S. Eliasson & E. H. Jahr (eds.), *Studies for Einar Haugen*, 343-365. Berlin: Mouton de Gruyter.

Pierrehumbert, J. & D. Talkin (1992). Lenition of /h/ and glottal stop. In G. J. Docherty & D. R. Ladd (eds.), *Papers in Laboratory Phonology II: gesture, segment, prosody*, 90-117. Cambridge: Cambridge University Press.

Prince, A. S. & P. Smolensky (1993). Optimality Theory: constraint interaction in generative grammar. *Technical report # 2 of the Rutgers Center for Cognitive Science*, Rutgers University.

Robins, R. H. (1970). Aspects of Prosodic Analysis. In F. R. Palmer (ed.), *Prosodic Analysis*, 188-200. Oxford: Oxford University Press.

Rothenberg, M. (1968). *The breath-stream dynamics of simple-released-plosive production*. Basel: Karger.

Schane, S. S. (1984a). The fundamentals of Particle Phonology. *Phonology Yearbook* 1. 129-156.

Schwartz, J. L., J. Robert-Ribes & P. Escudier (1998). Ten years after Summerfield: a taxonomy of models for audiovisual fusion in speech perception. In R. Campbell, B. Dodd & D. Burnham (eds.), *Hearing by eye, II: perspectives and directions in research on audiovisual aspects of language processing*, 85-108. Hove: Psychology Press.

Selkirk, E. O. (1982). The syllable. In H. van der Hulst & N. Smith (eds), *The structure of phonological representations,* Part II, 337-84. Dordrecht: Foris.

Steriade, D. (1993). Closure, release and nasal contours. In M. Huffman & R. Krakow (eds.), *Nasality*, 401-470. San Diego, CA: Academic Press.

Stevens, K. N. & S. E. Blumstein (1979). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64. 1358-1368.

Stevens, K. N. & S. E. Blumstein (1981). The search for invariant acoustic correlates of phonetic features. In P. D. Eimas & J. L. Miller (eds.), *Perspectives on the study of speech,* 1-35. Hillsdale, NJ: Erlbaum.

Traunmüller, H. (1994). Conventional, biological and environmental factors in speech communication: a modulation theory. *Phonetica* 51. 170-183.

Trubetzkoy, N. S. (1939). *Grundzüge der Phonologie. Travaux du cercle linguistique de Prague* 7.

Twaddell, W. F. (1935). On defining the phoneme. *Language Monograph* 16. Reprinted in M. Joos (ed.), 1958, *Readings in Linguistics* I, 55-80. New York.

Urua, E. E. (1990). Aspects of Ibibio phonology and morphology. PhD dissertation, University of Ibadan.

Urua, E. E. (2000). *Ibibio phonetics and phonology*. Cape Town: Centre for Advanced Studies of African Society.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science* 167. 392-393.

Warren, R. M. & C. J. Obusek (1971). Speech perception and phonemic restorations. *Perception and Psychophysics* 9. 358-362.

Westbury, J. & P. Keating (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145-166.

Williams, G. (1998a). A pattern recognition model for the phonetic interpretation of elements. *SOAS Working Papers in Linguistics and Phonetics* 7.

Williams, G. (1998b). The phonological basis of speech recognition. Unpublished PhD dissertation, SOAS, University of London.

Zoll, C. (1998). Positional asymmetries and licensing. Ms, MIT. ROA-282-0998.