Available online at www.sciencedirect.com

SCIENCE DIRECT®

Lingua 116 (2006) 1483–1494

Lingua

www.elsevier.com/locate/lingua

The phonology of being understood: Further arguments against sonority

John Harris

Department of Phonetics and Linguistics, University College London, Gower Street, London WC1E 6BT, United Kingdom

Received 16 March 2005; received in revised form 21 July 2005; accepted 28 July 2005

Abstract

Phonotactic restrictions reflect preferences in the way the carrier signal in speech is modulated from one moment to the next. This effect is largely unaccounted for in sonority-based analyses of phonotactics. The phonetic specification of sonority remains in any case controversial. However, treating it as an independently phonological entity is not an appealing option. Unlike distinctive features, sonority makes no contribution to the core of phonological knowledge that enables listener-talkers to attach linguistic meaning to modulations of the speech signal. Modulations carry the linguistic message, while the carrier enables the message to be heard. The sonority proposal attempts to characterise how messages are heard and has little to say about how they are understood.

Keywords: Phonotactic restrictions; Sonority; Modulation theory

1. Introduction

Languages typically impose quite severe restrictions on the ability of speech sounds to follow one another in phonological strings.[1] This patterned phonotactic behaviour has traditionally been attributed to sequencing restrictions imposed by SONORITY, loosely definable as relative loudness, perceptibility or acoustic intensity. More sonorous sounds, it is claimed, prefer to occur next to less sonorous sounds.

A well-established objection to sonority is that it lacks a consistent phonetic correlate. Nevertheless the concept continues to be widely employed in phonological theory. Wherever its

E-mail address: j.harris@ling.ucl.ac.uk.

[1] My thanks to the following for valuable comments on earlier drafts: Eric Carlson, Geoff Lindsey and two anonymous *Lingua* reviewers.

0024-3841/$ – see front matter
doi:10.1016/j.lingua.2005.07.009

use is explicitly defended, it is usually on the grounds that it represents some independent phonological entity that does not necessarily manifest itself in any phonetically homogeneous manner. Retreating to this position raises its own set of problems, some of which will be highlighted in this paper.

According sonority an autonomous phonological status invites comparison with another, much better established set of phonological entities designed to represent segmental quality, namely phonological features. Not only do these have clearly identifiable phonetic signatures, but they also fulfil the linguistic-symbolic function of signalling meaning contrasts. In the absence of any consistent phonetic basis, we would be doubly expectant to find sonority displaying at least some of the same distinctive functionality as features. The expectation is not borne out: sonority differences are never contrastive in the way that differences defined in terms of individual features can be.

This lack of distinctive function points up a more fundamental reason for doubting the phonological credentials of sonority. The reason emerges when we recall the standard definition of speech as a modulated carrier signal: the modulations bear the linguistic message, while the carrier enables the message to be heard. In focusing on acoustic intensity and loudness, the sonority proposal is essentially concerned with the nature of the carrier signal. That is, it is concerned primarily with the audibility of the linguistic message rather than with the message itself. This point is reflected in the fact that the fluctuation of intensity across utterances provides a poor guide to the flow of linguistic-phonetic information. For example, intervals of highest intensity – the sonority peaks associated with vowels – are typically the points of lowest contrastive potential. This is related to the fact that vowel systems are consistently smaller than consonant systems.

On one view, being dissociated from the linguistic message is enough to disqualify sonority as a component of phonological grammar. This is the conclusion we must draw if phonological grammars are conceived of as containing no more than the conventionalised knowledge that enables linguistic messages to be received and transmitted by means of speech sounds.

According to a broader view, the phonological grammar contains not just conventionalised knowledge but also what might be termed 'natural' knowledge – a tacit appreciation of the phonetic factors that shape the structure of conventionalised knowledge (for a survey of recent literature that adopts this standpoint, see Gordon, in press). On this view, a case might be made for sonority as a grammatical entity if it could be shown to explain how particular properties of the carrier signal might restrict the way one linguistically significant acoustic event follows another in speech.

There is a clear auditory-perceptual benefit to keeping one event as distinct as possible from the next. But the auditory-acoustic distance between two events cannot be reduced to differences in loudness or intensity, the supposed correlates of sonority. Acknowledgement of this point informs an alternative explanation of phonotactic restrictions, one that measures segment-by-segment change in terms of a set of acoustic parameters that includes not just intensity but also spectral shape, periodicity and fundamental frequency.

§2 reviews the problematic phonetic status of sonority and introduces several additional, phonological reasons for rejecting it. §3 outlines the proposal that phonotactic restrictions are driven by a need to optimise the auditory-perceptual distance between successive speech sounds. §4 presents evidence that the intensity envelope in speech is largely unrelated to the flow of linguistic-phonetic information across utterances. §5 considers some of the implications that dispensing with sonority has for feature theory.

2. Sonority

According to sonority theory, speech sounds possess a scalar property that allows them to be ranked on the following hierarchy, ranging from least to most sonorous: plosives $<$ fricati-fricatives $<$ nasals $<$ liquids $<$ glides/vowels. This is the ‘minimal’ hierarchy, one that is claimed to be valid for all languages (Clements, 1990; Baertsch, 2002). Refinements that define rankings among smaller segment classes usually come with the rider that they are subject to cross-linguistic variation.

The main use of the sonority hierarchy is to account for restrictions on sound sequences, particularly within and between syllables. A syllable is said to consist preferentially of a sonority peak (typically a vowel) flanked by consonants that decrease in sonority the further they lie from the peak. In the case of consonant clusters, sonority sequencing restrictions prefer a rising sonority slope in complex onsets (as in *pl, tr, tw*, etc.) and a falling slope in heterosyllabic sequences (as in *mp, ld, ft*, etc.).

Sonority has a long history in phonological theory (for early references, see for example Sievers, 1881; Jespersen, 1904; Saussure, 1916). It figures prominently in more recent work (see for example Selkirk, 1982; Clements, 1990) and continues to be extensively employed in Optimality Theory (see for example Prince and Smolensky, 1993; Zec, 1995; Gouskova, 2004).

The sonority hierarchy is initially established on the basis of cross-linguistic observation. At this point, like the inverse scale of phonological ‘strength’, it is no more than a taxonomic redescription of the patterns it is designed to account for (see Harris, 1985, chap. 2). Its main value is heuristic, aiding the formulation of generalisations about sound sequences that would perhaps otherwise have remained unexpressed or even undiscovered. For the hierarchy to take on explanatory value, it needs to be defined in terms of factors that are independently known to shape the design of phonological systems. There is a general consensus that the relevant factors are phonetic, involving some notion of articulatory aperture, acoustic intensity or auditory-perceptual salience. However, for a device with such an established place in phonological theory, a satisfactory phonetic definition of sonority in these terms has remained surprisingly elusive.

A specification of sonority based on articulatory aperture (as proposed by Saussure (1916) for example) can be straightforwardly applied to oral segments. If aperture is understood as degree of oral stricture, this yields a ranking corresponding to rising sonority for plosives, fricatives, liquids and vocoids. In order to accommodate the intermediate position of nasals on the hierarchy, the notion of aperture needs to be extended to include velic lowering. The combined contribution of the oral and nasal cavities to sonority might be defined aerodynamically, for example in terms of intraoral air pressure or total airflow. However, Parker’s (2002) experimental study of English and Spanish reveals that neither of these measures provides as good a fit with the sonority hierarchy as intensity.

Opening the oral cavity and venting the velic port might be viewed as independent articulatory means of producing a single acoustic effect of increasing intensity. In that case, output of acoustic energy would be expected to provide the most reliable physical measure of sonority (see Hankamer and Aissen, 1974). This is indeed what Parker’s (2002) study shows. It is certainly true that the high-sonority nuclei of syllables typically map to intensity peaks in utterances (Mermelstein, 1975). However, intensity measurements do not provide a consistent fit with the phonologist’s syllabifications. For one thing, not every local intensity maximum corresponds to a syllable nucleus. The method becomes particularly unreliable in the case of consonant clusters involving fine-grained local differences in intensity. This is especially true where strident fricatives are involved, since these can display higher intensity values than supposedly higher-sonority segments such as nasals.

It might be concluded that sonority does not map to any unitary physical property but is rather a cover term for a collection of independent acoustic properties that contribute to an overall dimension of perceptibility or auditory-perceptual salience (cf. Clements, 1990). Price (1980) suggests that three such components are involved: presence versus absence of a clear formant structure, presence versus absence of periodicity, and steady-state versus transient formant patterns. Each of these might be understood as contributing to overall acoustic intensity.

The main phonetic objection to sonority is not to do with whether it correlates in any consistent way with intensity (at least under certain experimental conditions it can be seen to do so). Rather it has to do with the assumption that intensity determines the role played by perceptibility in phonotactics. As we will see below, perceptibility is based on the extent to which a sound stands out from the carrier signal and from neighbouring sounds, and this involves more acoustic parameters than intensity alone (see especially Ohala, 1992).

Aside from any question marks over its phonetic status, sonority is in any event potentially problematic in at least three other respects.

One problem concerns the representational status of sonority. At least with standard SPE-derived features, the sonority value of a segment cannot be directly read off phonological representations but has to be calculated by reference to an external look-up table – the sonority hierarchy. At the very least this is unparsimonious. An obvious solution is to adapt feature theory in such a way that sonority specifications can be directly built into representations. According to one such proposal, a multivalued sonority feature replaces major-class features (see for example Hankamer and Aissen, 1974; Selkirk, 1982; Parker, 2002). Another approach is to redefine familiar features in such a way that a segment's sonority is directly reflected in the number of plus values it bears (Clements, 1990) or in its specificational complexity (Rice, 1992).

Another issue concerns the breadth of sonority's empirical coverage. One of the main advantages of the sonority hierarchy is the facility with which it allows us to make broader generalisations than are possible with individual features. However, its coverage is arguably not general enough. The hierarchy orders sound classes primarily on the basis of their manner characteristics. There is a substantial set of restrictions on cluster phonotactics that lie beyond the reach of sonority, even though they occur in exactly the same environments as restrictions that can be described in terms of sonority sequencing. Many of these have to do with place contrasts. Well-known examples include the homorganic sequences *tl/dl/θl* and *pw/bw/fw*. Given their rising sonority slope, these should make good onset clusters. However, they are at best highly marked (they are impossible in English, for example). The usual response has been to assume that these exceptions are exclusively due to independent place restrictions and have nothing to do with sonority at all (see Selkirk, 1982, for discussion). That may be, but a sonority account would have to give way to any alternative that managed to subsume these place-based restrictions under some more general analysis of cluster phonotactics. This is exactly what an approach based on inter-segment perceptual distance succeeds in doing, as we will see below.

The stop-liquid examples illustrate another shortcoming of sonority. Cross-linguistically, while *tl/dl* are disfavoured as complex onsets, the reverse sequences *lt/ld* are favoured as coda-onset clusters (as in English *filter*, *boulder*). Sonority sequencing restrictions prefer a sonority slope in both contexts, rising in onsets and falling in syllable contact. All else being equal, the slopes would be expected to be mirror images of one another. But this situation is rarely encountered in individual languages. It is much more usual to find that the sonority gradient is required to be steeper in onsets than heterosyllabically – what Clements (1990) calls the 'sonority cycle'.

Take two-consonant clusters containing a plosive in English, for example. If the plosive is the first member of a complex onset, the second consonant has to be a high-sonority liquid or glide (e.g. *play*, *twelve*). In contrast, if the plosive is preceded by a coda, this can contain any manner of consonant, ranging from a high-sonority liquid (e.g. *filter*) to a low-sonority stop (e.g. *chapter*). Asymmetries of this sort can be captured by imposing different sonority distances between neighbouring consonants in the two syllabic contexts (cf. Selkirk, 1982; Clements, 1990). It should go without saying, however, that this solution is purely stipulative.

3. Perceptual distance

The notions of sonority sequencing and sonority distance are surely on the right track in at least one respect: they tap into the basic insight that sounds in a sequence prefer to be different from one another (on this point, see Steriade, 1982). Intuitively at least, this would seem to be advantageous from an auditory-perceptual viewpoint. However, the perceptual distance between neighbouring sounds is determined not on the basis of raw intensity but on the basis of a suite of different acoustic properties. To get a handle on this point, let us return to the definition of speech as a modulated carrier signal.

In the usual case, an unmodulated carrier signal has the quality of schwa – the sound produced by a neutrally open vocal tract. Acoustically, it lacks spectral prominences and is typically periodic. TYPICALLY is important here, because periodicity is not a necessary component of the carrier. To fulfil its main communicative function of making linguistic messages audible, the carrier requires a sustained sound source. For the average healthy human talker, the most effective way to create this is by means of vocal-fold vibration driven by outgoing air from the lungs. Alternative methods of producing a continuous sound source can be employed, including whisper and oesophageal voice.[2] Not all of these methods create a periodic waveform. This point becomes important when it comes to deciding whether and how periodicity should be phonologically encoded (on which more below).

The carrier signal bears information that is non-linguistic. It can reveal details about the talker's identity, state of mind and location (see Traunmüller, 1994). From the carrier, the listener can make inferences about such questions as whether the talker is human (as opposed to a minah bird, for example), how large they are, what sex they are, how old they are, whether they are angry, whether they're chewing gum, how far away they are.

Acoustic events that modulate the carrier signal in speech bear information that is specifically linguistic in the sense of providing auditory-acoustic cues that aid lexical access and morphosyntactic parsing.

It is evident that humans have a facility for 'demodulating' the speech signal (Traunmüller, 1994) – of separating linguistic information from the non-linguistic. The validity of this observation is not undermined by the fact that we do not yet fully understand how this filtering process works. The most obvious indication that the process does occur lies in the phenomenon of normalisation, whereby listeners are able to perceive the 'same' phonetic quality in signals produced by different speakers with vocal tracts of radically different sizes (see Klatt, 1986). This

[2] Other mechanisms can be exploited short-term for linguistic-contrastive purposes, such as temporarily switching the direction of airflow or engaging other airstream initiators.

behaviour is presumably related the fact that speech perception operates in distinct modes according to whether the relevant percepts are linguistically meaningful or not. Linguistically contrastive acoustic differences are perceived categorically (Liberman, 1970; Mattingly et al., 1971; Jusczyk, 1986), while those bearing non-linguistic information can be perceived in a more or less continuous fashion. Take the example of vowel quality. On the one hand, linguistically significant vowel contrasts are perceived in terms of relatively gross categories. Each of these acts as a perceptual 'magnet', blunting the acuity with which listeners are able to discriminate formant differences that lie near the centre of the category (Kuhl, 1987). The grossness of these categories is reflected in the relatively small size of languages' vowel systems. On the other hand, social and personal evaluations of talkers can be made on the basis of formant differences that are often too small ever to be harnessed for contrastive purposes (see for example Labov et al.'s, 1972 measurements of socially and regionally stratified vowel variation in English).

How much of the variegated information present in the speech signal is coded in phonological grammars? Under a narrow conception of the phonological grammar as a component of language competence, the answer is quite clear: only the information that is specifically linguistic. Any property attributable to the carrier signal, being linguistically insignificant, should not be phonologically encoded. In particular, it does not warrant specification in terms of phonological features.

Under the broader view of phonological grammar as also including natural knowledge, properties of the carrier signal would merit phonological representation if they could be shown to influence the structure of linguistically significant phonetic information. But even in this broader model, it is necessary to reserve a place for a 'core' of phonological knowledge dedicated to the interpretation of specifically linguistic information, in light of the distinct auditory-perceptual behaviour it evinces. (Whether this core is represented in terms of some independent sub-module or is distributed across the phonological grammar is a separate issue.)

A consideration of the features standardly used to define sonority confirms that it does not satisfy the core criterion of linguistic significance. Take for example the combination of major-class values that identifies vowels as the most sonorous group of sounds: [−consonantal], [+sonorant] and [+continuant]. All of these specifications describe aspects of the carrier signal – either its acoustic characteristics (periodicity) or the articulatory and aerodynamic mechanisms necessary to produce it (unobstructed vocal tract, absence of a build-up of air pressure). The lack of linguistic value associated with these properties is underlined by the fact that they are universally redundant; that is, they are not exploited for contrastive purposes in vowel systems.

Linguistically significant information in vowels is borne by signal modulations that take the form of spectral prominences (due to the convergence of formants) which deviate from the carrier's neutral resonance baseline. These modulations bear information about contrasts in vowel quality and thus merit representation in the grammar. Whether this is couched in terms of familiar features such as [back], [high] and [round] or some other classification is a separate issue (see Harris and Lindsey, 2000). Note that these quality specifications play little or no role in the definition of sonority (other than in distinguishing relatively more open, more sonorous vowels from closer, less sonorous vowels).

In short, sonority has no place in the core of phonological knowledge that enables listeners to attach linguistic significance to the modulations they perceive in speech signals. According to the view of the phonological grammar as encompassing natural knowledge, there is still the possibility that sonority at least deserves grammatical representation on the grounds that it influences the way one modulation follows another. In particular, it might be claimed that sonority identifies an aspect of the carrier signal in terms of which preferred differences between sounds in sequence are defined. However, there are good grounds for rejecting the notion that these preferences are reducible to differences in acoustic intensity.

The main parameters along which linguistically significant acoustic events can modify the carrier signal in speech include not just intensity but also spectral shape, periodicity and fundamental frequency (see Ohala, 1992, for extensive discussion and references). The magnitude of a modulation can be measured in terms of the trajectory it follows through a multidimensional acoustic space defined by these parameters. The longer the trajectory, the more perceptually salient a modulation is.

The relative salience of a given sound can be gauged along two axes, based on the extent to which the modulation with which it is associated deviates (i) from the carrier signal and (ii) from the sound(s) next to it. These two aspects can be illustrated by the idealised sound sequences *əpə* and *əptə*. In both cases, the vowels can be taken as establishing a carrier baseline from which we can evaluate the salience of the intervening consonants.

In *əpə*, the modulation associated with the consonant traces an extended acoustic path that traverses all of the main parameters. There is an abrupt and sustained drop in amplitude, corresponding to the closure phase of the stop. There is an F0 discontinuity, caused by the cessation of voicing during the consonant. There is a burst of aperiodic energy on the release of the plosive. There are changes in spectral shape that help cue the place category of the stop – formant transitions during the approach and release phases and the frequency characteristics of the noise burst. Given the schwa quality of the flanking vowels, these events can primarily be viewed as acoustic landmarks that modulate the carrier signal (see Stevens, 2002). In this context, they combine to provide a robust set of auditory-acoustic cues to the identity of the *p* (Wright, 2004).

Now consider *əptə*. The approach phase of the *p* and the release phase of the *t* display similar acoustic properties to the corresponding phases of *p* in *əpə*. However, the juxtaposition of two stops restricts the extent to which the speech signal can be modulated from one consonant to the next. In two-stop clusters, it is quite common to find that the closure phase of the second overlaps with the release phase of the first (cf. Browman and Goldstein, 1992). In that case, some of the acoustic events otherwise associated with such consonants fail to occur, in particular the release burst on *p* and the formant transitions at the approach to *t*. A potentially valuable set of cues to the place identity of the stops is thereby obscured (again see Wright, 2004). (For reasons to be spelt out below, the masking effect is potentially more damaging to the detectability of the first stop than to the second.) Compare this with *əstə*, where the cues to the first consonant remain largely unaffected by the presence of the second. The manner category of the fricative is cued by steady-state aperiodicity, the frequency characteristics of which provide a robust cue to its place category.

Measured both against the background carrier and against neighbouring sounds, the magnitude of the modulation associated with *p* is greater in *əpə* than in *əptə*. In consequence, *p* is more salient in the first context, in the sense of being more easily detectable by the listener. This particular comparison illustrates the more general point that phonotactic restrictions are evidently related to the varying facility with which the listener is able to extract cues to the identity of segments from the acoustic signal in different phonological positions. In the *əpə*/*əptə* examples, the 'cueing potential' of the intervocalic context is greater than the preconsonantal context (Steriade, 1997, 1999).

Cross-linguistic observation confirms that clusters of oral stops are more marked than fricative-stop or sonorant-stop clusters (Kawasaki, 1982). Under a sonority account, this is attributed to a preference for a greater sonority distance between the two members of a cluster.

Under an account based on cueing potential, the markedness asymmetry has to do with differing perceptual distances based on the collection of auditory-acoustic parameters mentioned above (Ohala, 1992; Steriade, 1999). Information about the identity of the first of two contiguous consonants is more readily detectable when that segment is distinct in manner from the second (as in V*st*V) than when it is not (as in V*pt*V).

The two notions of distance are not at all equivalent. For some types of cluster, including the fricative-stop and stop-stop examples just mentioned, the two approaches do converge. However, there are many types of restrictions on segment sequencing that can be explained in terms of perceptual distance but simply lie outside the ambit of sonority. These include the examples mentioned in the previous section, to which we now return.

Let us reconsider the question of why tautosyllabic and heterosyllabic consonant clusters are not mirror images of one another. Take the example of the least marked onset, one containing a plosive. The contrastive potential of any position following it within the same onset is typically more tightly constrained than that of a preceding coda. Under a sonority account, recall, this imbalance has to be stipulated in terms of differing slope requirements. Under an account based on perceptual distance, the difference between the two types of cluster is attributable to an on-off asymmetry in the way the auditory nerve responds to incoming acoustic signals (Bladon, 1986). Spectral changes evoke different responses according to whether they produce an onset or an offset of neural firing. The former are much more perceptually salient than the latter. In a CV context, the transition from the consonant is primarily a spectral onset. In a VC context, the transition into the consonant is primarily a spectral offset. Amongst other things, this explains why release-phase cues to the identity of a plosive are perceptually more important than those in the approach phase (Wright, 2004). Any consonant appearing next to the plosive will potentially interfere with these cues. However, from the viewpoint of perceptibility, it is worse to obscure the spectral onset of the plosive than its spectral offset. It is thus less optimal to have a consonant following the plosive than to have one preceding it. Moreover, if a following consonant is to be there at all, it is better for it to be a resonant rather than, say, another obstruent. The spectral characteristics of the plosive's release burst can be superimposed on a resonant but are liable to be masked by the closure phase of a following stop or distorted by the noise spectrum of a following fricative.

This brings us once again to the question of why there is a cross-linguistic dispreference for onset clusters composed of *pw/bw* or *tl/dl*. The basic reason seems to be that the members of each of these pairs are perceptually too similar (Kawasaki, 1982; Ohala, 1992). In the case of the labial clusters, the spectrum for *w* (low F1 and F2) is not substantially different from that following the release of a labial stop. The segment-by-segment modulation of spectral shape is insufficiently large to support a reliable contrast with other stop+*w* clusters. An additional factor is possibly in play with *tl/dl*. Here the reduction in amplitude created by the lateral's medial oral stricture is likely to attenuate spectral cues to the place of the stop, thereby impairing the detectability of contrasts with other stop+*l* clusters (in particular *kl/gl*). The dispreference for *tl/dl* clusters is evidently not attributable to some notion of articulatory difficulty (related perhaps to homorganicity). The same phonetic sequences occur relatively frequently in the form of lateral affricates – that is, as single-position contour segments in which affricate homorganicity precludes the potential for place contrasts found in two-position clusters.

4. The information cycle

In view of its implicit focus on the shape of the carrier signal, it is not surprising that the sonority literature has little to say about the patterning of linguistically significant information in the speech signal. Nevertheless, it is instructive to reflect on how sonority theory relates to this issue.

There is an inherent property of speech that seems at first sight paradoxical: most of the sound energy is concentrated in vowels, while most of the linguistically significant information is concentrated in consonants. Very crudely speaking, this means that the 'sonority cycle' is inversely phased in relation to a cycle of linguistic information. Points of sustained high intensity (vowels) modulate the carrier signal to a lesser extent than intervening points (consonants) which, though of lower intensity, involve relatively greater and more rapid spectral changes. Having available a more extensive range of modulations bestows on energy troughs a greater potential to bear linguistic information than peaks. Let us briefly consider a range of facts that confirm this point.

There is a strong tendency for consonant systems to be larger than vowel systems. Among the 317 languages in the UPSID database, the mean inventory size is 22.8 for consonants and 8.7 for vowels (Maddieson, 1984). The greater contrastive potential of consonants can reasonably be taken as symptomatic of a greater capacity to bear linguistic information in speech signals.

Altering or obscuring vowel quality has relatively little impact on the intelligibility of running speech. This effect is observed in experimental studies where listeners successfully interpret utterances in which all vowels have been instrumentally replaced by a single quality such as schwa. Performing a corresponding substitution on consonant quality (for example by using *h* as the replacement segment) significantly impairs intelligibility (see for example Carlson and Granström, 1977). The vowel effect can also be exploited in language games, such as the singing tradition in which all the vowels in a verse of song are replaced by a single quality that is then varied on successive repetitions of the verse.

In alphabetic writing, the omission of vowels impairs reading intelligibility significantly less than the omission of consonants (compare I O EIE with SMTH FR PRSDNT). This principle is incorporated into speed writing, including shorthand and phone text messaging.

In some languages, lexical-category morphemes are represented by discontinuous segmental melodies. To the best of my knowledge, these are always composed of consonants, as in the well-known example of Semitic. The vowels that break these melodies up represent either affixes or some non-contrastive quality that is epenthesised to satisfy the demands of syllabification. In languages of this type, consonants thus bear the heavy functional load of contrasting open-class items (nouns, verbs, adjectives), while vowels have the much lighter responsibility of contrasting closed-class morphemes.

Dialect differences by and large show greater variation in vowel systems than in consonant systems. This point is quite dramatically illustrated by the vowel systems of Germanic languages, where historical shifts can give rise to radically different qualities within a given set of cognate words, even in closely related dialects. Take the example of English historically short *a* (as in *sat*, *man*, *pass*). In many dialects, this vowel has undergone lengthening in certain phonological contexts. This has triggered a series of qualitative changes that have produced reflexes as far apart as high front and low back. So, for example, we find *ɪə* in *bad*, *man*, *pass* in parts of the northeastern United States versus *ɑː* in *pass*, *bath*, *dance* in southern England and the southern hemisphere. Nothing of this order of phonetic difference occurs in the consonant system of modern English. Evidently much of the information borne by vowel quality is indexical, divulging details about speakers' social and regional backgrounds. Historical vowel shifts can give rise to a range of indexically differentiated variants without seriously interfering with the communication of specifically linguistic information. Consonant changes are not afforded the same degree of leeway, suggesting that they are constrained by having to bear a heavier linguistic-functional load than vowels.

In general, the informational yield of energy peaks in speech is high in personal (and affective) quality and relatively low in linguistic-phonetic quality. Conversely, energy troughs are rich in linguistic-phonetic information and relatively poor in personal information. In other words, vowels reveal much about who you are or how you feel, while consonants reveal more about what you are saying.

5. Implications for feature theory

Sonority has no place in core phonological knowledge on the grounds that it does not link to the specifically linguistic content of speech signals. Applying the same criterion to phonological form in general has far-reaching consequences for feature theory.

An immediate consequence is that the core can only accommodate feature specifications which map exclusively to modulations of the carrier signal. There is no place for specifications which map to properties of the carrier itself (such as the orthodox specification of vowels and glides as [−consonantal, +sonorant, +voice] mentioned above).

On the other hand, the phonetic expression of each feature in the core is definable in relation to the carrier – specifically, in terms of how the acoustic event with which the feature is associated perturbs the baseline presented by the carrier. Pursuing this line of thinking, we reach a point where we can dispense with feature redundancy. The phonetic signature of any given feature can be expressed in isolation from other features, because it will always be supported by the carrier signal (for more discussion of this point, see Harris and Lindsey, 1995). For example, the phonetic signature of the vowel *a* modulates the evenly spaced formant structure of the carrier by compressing the first two formants into a single spectral peak occupying the middle of the frequency range relevant to vowel quality.

This in turn prompts a rethink of the notion that redundant feature values can enhance the phonetic expression of distinctive values (Stevens and Keyser, 1989). Feature enhancement is standardly understood as a grammar-internal relation, in which one phonological category supports another. For example, [+voice] is said to enhance features such as [back], [high] and [round] in the specification of vowel quality. Rethought, the relation becomes one in which a non-core property associated with the carrier signal supports a core category associated with a modulation of the carrier. In the example just given, periodicity in vowels is a non-core property that enhances a particular core feature of vowel quality. According to this interpretation, enhancement is essentially about supporting the audibility of a phonological feature.

6. Conclusion

Perceptibility in speech is based on the principle that change is more salient than stability (again, see Ohala, 1992). This is what drives restrictions on sound sequences. There are several fundamental questions we can ask about this principle. Does it take a specific form that is peculiar to speech? Is it innate? Is it part of a listener-talker's linguistic competence?

On this last question, there is no a priori reason to suppose that the principle of change is integrated into phonological competence, as has often been claimed for sonority. That is, acknowledging the importance of the principle in shaping phonological systems does not commit us to the position that it is represented in the grammar, for example in the form of an Obligatory Contour Principle or as a set of functionally live constraints. It is one thing to note that listener-talkers must have tacit knowledge of restrictions on sound sequences (which, among other things, establish conventions governing possible word forms in their language). It is quite another to claim that they also have knowledge of the forces responsible for these restrictions (although for explicit arguments in favour of this general view, see Hayes, 2004). Under the simplest view of phonological grammars as containing exclusively linguistic knowledge, the principle of salience as change can be seen as a purely grammar-external force that guides the evolution of sound systems over time (cf. Ohala, 1995; Blevins, 2004; Lu, 2004).

What is clear is that the perceptual distance between neighbouring sounds in speech involves more than differences in acoustic intensity, the supposed correlate of sonority. The energy envelope in speech forms part of a carrier signal for the linguistic message, but it is not part of the message itself.

We are continually reminded of the axiom that 'we speak to be heard to be understood' (Jakobson et al., 1952). Perceptual distance is relevant to being understood, while the energy envelope is relevant to being heard. The sonority proposal is irrelevant to the first of these and a flawed model of the second.

References

Baertsch, K., 2002. An optimality-theoretic approach to syllable structure: the split margin hierarchy. Ph.D. dissertation, Indiana University.

Bladon, A., 1986. Phonetics for hearers. In: McGregor, G. (Ed.), Language for Hearers. Pergamon, Oxford, pp. 1–14.

Blevins, J., 2004. Evolutionary Phonology: The Emergence of Sound Patterns. Cambridge University Press, Cambridge.

Browman, C.P., Goldstein, L., 1992. Articulatory phonology: an overview. Phonetica 49, 155–180.

Carlson, R., Granström, B., 1977. Perception and synthesis of speech. Speech Transmission Laboratory Quarterly Progress and Status Report 1, 1–16.

Clements, G.N., 1990. The role of the sonority cycle in core syllabification. In: Kingston, J., Beckman, M.E. (Eds.), Papers in Laboratory Phonology I; Between the Grammar and Physics of Speech. Cambridge University Press, Cambridge, pp. 283–333.

Gordon, M., in press. Functionalist phonology. In: de Lacy, P. (Ed.), The Cambridge Handbook of Phonology. Cambridge University Press, Cambridge.

Gouskova, M., 2004. Relational hierarchies in Optimality Theory: the case of syllable contact. Phonology 21, 201–250.

Hankamer, J., Aissen, J., 1974. The sonority hierarchy. In: Bruck, A., Fox, R.A., La Galy, M.W. (Eds.), Papers from the Parasession on Natural Phonology. Chicago Linguistic Society, Chicago, pp. 131–145.

Harris, J., 1985. Phonological Variation and Change. Cambridge University Press, Cambridge.

Harris, J., Lindsey, G., 1995. The elements of phonological representation. In: Durand, J., Katamba, F. (Eds.), Frontiers of Phonology: Atoms, Structures, Derivations. Longman, Harlow, Essex, pp. 34–79.

Harris, J., Lindsey, G., 2000. Vowel patterns in mind and sound. In: Burton-Roberts, N., Carr, P., Docherty, G. (Eds.), Phonological Knowledge: Conceptual and Empirical Issues. Oxford University Press, Oxford, pp. 185–205.

Hayes, B., 2004. Phonetically-driven phonology: the role of Optimality Theory and inductive grounding. In: McCarthy, J.J. (Ed.), Optimality Theory in Phonology: A Reader. Blackwell, Oxford, pp. 290–309.

Jakobson, R., Fant, G., Halle, M., 1952. Preliminaries to Speech Analysis. MIT Press, Cambridge, MA.

Jespersen, O., 1904. Phonetische Grundfragen. Teubner, Leipzig.

Jusczyk, P.W., 1986. A review of speech perception research. In: Kaufman, L., Thomas, J., Boff, K. (Eds.), Handbook of Perception and Human Performance. Wiley, New York, pp. 27–57.

Kawasaki, H., 1982. An acoustical basis for universal constraints on sound sequences. Ph.D. dissertation, University of California, Berkeley.

Klatt, D.H., 1986. The problem of variability in speech recognition and in models of speech perception. In: Perkell, J.S., Klatt, D.H. (Eds.), Invariance and Variability of Speech Processes. Erlbaum, Hillsdale, NJ, pp. 300–319.

Kuhl, P.K., 1987. Perception of speech and sound in early infancy. In: Salapatek, P., Cohen, L. (Eds.), Handbook of Infant Perception, vol. 2. Academic Press, New York, pp. 275–382.

Labov, W., Yaeger, M., Steiner, R., 1972. A Quantitative Study of Sound Change in Progress. US Regional Survey, Philadelphia, PA.

Liberman, A.M., 1970. Some characteristics of perception in the speech mode. Perception and its Disorders 48, 238–254.

Lu, A.C.L., 2004. Explaining final obstruent voicing in Lezgian: phonetics and history. Language 80, 73–97.

Maddieson, I., 1984. Patterns of Sounds. Cambridge University Press, Cambridge.

Mattingly, I.G., Liberman, A., Syrdal, A.K., Halwes, T., 1971. Discrimination in speech and nonspeech modes. Cognitive Psychology 2, 131–157.

Mermelstein, P., 1975. Automatic segmentation of speech into syllabic units. Journal of the Acoustical Society of America 58, 880–883.

Ohala, J.J., 1992. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. CLS: Papers from the Parasession on the Syllable 319–338.

Ohala, J.J., 1995. Phonetic explanations for sound patterns: Implications for grammars of competence. In: Elenius, K., Branderud, P. (Eds.), Proceedings of the 13th International Congress of Phonetic Sciences, August 1995, vol. 2, Stockholm, pp. 52–59.

Parker, S., 2002. Quantifying the sonority hierarchy. Ph.D. dissertation, University of Massachusetts.

Price, P.J., 1980. Sonority and syllabicity: acoustic correlates and perception. Phonetica 37, 327–343.

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

Rice, K.D., 1992. On deriving sonority: a structural account of sonority relationships. Phonology 9, 61–100.

de Saussure, F., 1916. Cours de Linguistique Générale. Payot, Paris.

Selkirk, E.O., 1982. The syllable. In: van der Hulst, H.,Smith, N. (Eds.), The Structure of Phonological Representations, Part II. Foris, Dordrecht, pp. 337–384.

Sievers, E., 1881. Grundzüge der Phonetik. Breitkopf & Hartel, Leipzig.

Steriade, D., 1982. Greek prosodies and the nature of syllabification. Ph.D. dissertation, MIT.

Steriade, D., 1997. Phonetics in phonology: The case of laryngeal neutralization. Ms, UCLA.

Steriade, D., 1999. Alternatives to the syllabic interpretation of consonantal phonotactics. In: Fujimura, O., Joseph, B.D., Palek, B. (Eds.), Proceedings of the 1998 Linguistics and Phonetics Conference, Karolinum Press, Prague, pp. 205–242.

Stevens, K., 2002. Toward a model of lexical access based on acoustic landmarks and distinctive features. Journal of the Acoustical Society of America 111, 1872–1891.

Stevens, K., Keyser, S.J., 1989. Primary features and their enhancements in consonants. Language 65, 81–106.

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

Wright, R., 2004. A review of perceptual cues and cues robustness. In: Hayes, B., Kirchner, R., Steriade, D. (Eds.), Phonetically Based Phonology. Cambridge University Press, Cambridge, pp. 34–57.

Zec, D., 1995. Sonority constraints on syllable structure. Phonology 12, 85–129.