There is no level of phonetic representation[*]

JOHN HARRIS & GEOFF LINDSEY[**]

1 Introduction

What size are the primes of which phonological segments are composed? From the standpoint of orthodox feature theory, the answer is that each prime is small enough to fit inside a segment, and not big enough to be phonetically realised without support from other primes. Thus, [+high], for example, is only realisable when combined with values of various other features, including for instance [-back, -round, -consonantal, +sonorant] (in which case it contributes to the definition of a palatal approximant). This view retains from earlier phoneme theory the assumption that the segment is the smallest representational unit capable of independent phonetic interpretation.

In this paper, we discuss a fundamentally different conception of segmental content, one that views primes as small enough to fit inside segments, yet still big enough to remain independently interpretable. The idea is thus that the subsegmental status of a phonological prime does not necessarily preclude it from enjoying stand-alone phonetic interpretability. It is perfectly possible to conceive of primes as having autonomous phonetic identities which they can display without requiring support from other primes. This approach implies recognition of 'primitive' segments, each of which contains but one prime and thus reveals that prime's autonomous phonetic signature. Segments which are non-primitive in this sense then represent compounds of such primes.

This notion — call it the **autonomous interpretation** hypothesis — lies at the heart of the traditional notion that mid vowels may be considered amalgamations of high and low vowels. Adaptations of this idea are to found in the work of, among others, Anderson & Jones (1974) and Donegan (1978). Anderson & Jones' specific proposal is that the canonical five-vowel system should be treated in terms of various combinations of three primes which we label here [A], [I] and [U]. Individually these manifest themselves as the primitive vowels *a*, *i* and *u*

[*]This paper, a draft excerpt from Harris & Lindsey (forthcoming), was delivered to the Cognitive Phonology Workshop, University of Manchester, April 1993. Our thanks to Outi Bat-El, Wiebke Brockhaus, Phil Carr, Jacques Durand, Edmund Gussmann, Doug Pullyeblank and Neil Smith for their helpful comments on earlier drafts.

[**]Geoff Lindsey is at the University of Edinburgh. Correspondence address: Department of Linguistics, University of Edinburgh, George Square, Edinburgh EH8 9JX, Scotland (geoff@ling.ed.ac.uk).

respectively (see (1)a). As shown in (1)b, mid vowels are derived by compounding [A] with [I] or [U].

(1)

(a)	[A]	a	(b)	[A,I]	e
	[I]	i		[A,U]	o
	[U]	u			

Not all work which incorporates this type of analysis has maintained the notion of autonomous interpretation. In current Dependency Phonology, the direct descendant of Anderson & Jones' (1974) proposal, the view is in fact abandoned in favour of one in which primes such as those in (1) ('components') define only the resonance characteristics of a segment and must be supplemented by primes of a different sort ('gestures') which specify manner and major-class properties (see Anderson & Durand 1987, Anderson & Ewen 1987 and the references therein). A similar line has been followed in van der Hulst's 'extended' Dependency approach (1989, in press).

Nevertheless, the principle of autonomous interpretation continues to figure with varying degrees of explicitness in other approaches which employ the primes in (1). This is true, for example, of Particle Phonology (Schane 1984) and Government Phonology (Kaye, Lowenstamm & Vergnaud 1985, 1990), as well as of the work of, among others, Rennison (1984, 1990), Goldsmith (1985) and van der Hulst & Smith (1985). Although the use of the term element to describe such primes is perhaps most usually associated with Government Phonology, we will take the liberty of applying it generically to any conception of subsegmental content which incorporates the notion of autonomous interpretation. The analogy with physical matter seems apt in view of the idea that phonological elements may occur singly or in compounds with other elements. Moreover, to the best of our knowledge, Government Phonology is the only framework to have explicitly pushed the autonomous interpretation hypothesis to its logical conclusion — namely that each and every subsegmental prime, not just those involved in the representation of vocalic contrasts, is independently interpretable.

In this paper, we explore some of the consequences of adopting a full-blooded version of element theory. This entails a view of phonological derivation which is radically different from that associated with other current frameworks, particularly those employing feature underspecification. One significant implication of the autonomous interpretation hypothesis is that phonological representations are characterised by full phonetic interpretability at all levels of derivation. This in turn implies that there is no level of systematic phonetic representation. Viewed in these terms, the phonological component is not a device for converting abstract lexical

representations into ever more physical phonetic representations, as assumed say in underspecification theory. Instead, it has a purely generative function, in the technical sense that it defines the grammaticality of phonological structures. In the following pages, we will argue that these consequences of the autonomous interpretation hypothesis inform a view of phonology and phonetic interpretability that is fully congruent with current thinking on the modular structure of generative grammar.

The various element-based and related approaches share a further theoretical trait which distinguishes them from traditional feature theory: they all subscribe to the assumption that phonological oppositions are, in Trubetzkoyan terms, **privative**. That is, elements are single-valued (monovalent) objects which are either present in a segment or absent from it. Traditional feature theory, in contrast, is based on the notion of **equipollence**; that is, the terms of an opposition are assigned opposite and in principle equally weighted values of a bivalent feature. Of the various facets of element-based theory, monovalency is the one that has attracted the closest critical attention. The empirical differences between privative and equipollent formats continue to be disputed. Since the arguments have been well aired elsewhere (see especially van der Hulst 1989, den Dikken & van der Hulst 1989), we will simply side-step this issue here.

Our intention is to shift the focus of the comparison between features and elements to the fundamental conceptual differences that centre on the issue of interpretational autonomy, a matter that up to now has received only scant attention in the literature. In the next section, we discuss the place of the autonomy hypothesis in a view of phonology which at first blush may seem paradoxical but which is in fact fully in tune with a modular theory of language. On the one hand, the view is 'concrete' to the extent that phonological representations are assumed to be phonetically interpretable at all levels of derivation. On the other, the representations are uniformly cognitive; the primes they contain, unlike orthodox features, do not recapitulate modality-specific details relating to vocal or auditory anatomy. In §3, we go on to show how this view is implemented in the specification of elements involved in the representation of vowels.

2 Phonetic interpretation in generative grammar

2.1 Phonetic representation has no theoretical status

Recognising the interpretational autonomy of phonological primes gives rise to a view of phonological representations and derivations which is fundamentally different from that associated with orthodox features. Within the latter tradition, it

has long been usual to suppose that only a subset of the feature specifications appearing in final representation is present at the outset of derivation (Halle 1959). Underlyingly absent values, those that are non-distinctive or predictable, are filled in during the course of derivation by redundancy rules or phonological rules proper. This view achieves apotheosis in Radical Underspecification Theory, in which all predictable feature values are stripped from underlying representation (Archangeli 1984, 1988, Pulleyblank 1986). There has been a slight retreat from this extreme position in recent descendants of the theory. In the Combinatorial Specification approach of Archangeli & Pulleyblank, for example, a criterion of representational simplicity which favours maximal despecification of lexical representations may be overridden in certain cases where otherwise non-distinctive values can be shown to play an active role in underlying association patterns (1992: 88-89).

Coupled to the assumption that features lack interpretational autonomy, the classic underspecification arrangement implies that any non-final representation containing blank feature values is phonetically uninterpretable. The realisation of lexically represented values is contingent on the support of non-distinctive or predictable values; and the full complement of mutually supporting feature values is not mustered until the final stage of derivation, the level of systematic phonetic representation. This view too has been modified somewhat in more recent underspecification approaches. As a result of work by Keating (1988), Pierrehumbert & Beckman (1988) and others, it is now assumed that fragments of underspecification may persist into phonetic implementation (Archangeli & Pulleyblank 1992: 43, Pulleyblank, in press). This situation supposedly arises in certain instances of what used to be termed **sequence redundancy**, where the phonetic interpretation of a segment with respect to a particular feature is entirely predictable on the basis of the value this feature has in neighbouring segments. In a sub-class of such cases, it has been argued, the intervening sound fails to be assigned a value for the feature in question and is thus submitted to motor planning without an inherent articulatory target. The relevant articulatory dimension then allegedly manifests itself through simple linear interpolation between the motor targets associated with the specified values of the flanking segments. (Reservations about the allegedly targetless nature of such segments have been expressed by, among others, Boyce, Krakow & Bell-Berti 1991.) Even when adapted to allow for sporadic instances of persistent underspecification, it nevertheless remains true of this overall approach that a significant proportion of feature values, particularly those which are predictable from other values within the same segment (**segment redundancy**), must be specified in phonetic representation before phonetic interpretation is possible.

According to this line of thinking, one of the main jobs of the phonological rule component is to transform abstract representational objects into ever more

physical ones. This mismatch between underlying and surface representations is sometimes justified on the grounds that the two levels allegedly perform quite different functions: underlying representations, it is sometimes claimed, serve the function of memory and lexical storage, while surface representations serve as input to articulation and perception (Bromberger & Halle 1988). The validity or otherwise of this view hinges on a number of considerations, only some of which we have space to consider here.

One issue concerns the degree of circularity that is inherent in the strategy of stripping lexical representations down to the distinctive bone, a problem fully discussed by Mohanan (1991). In many instances, the presence of a given pair of phonological properties in a representation may involve balanced mutual dependence. That is, in such cases there is no non-arbitrary way of determining the directionality of the relation whereby one property is deemed lexically distinctive and the other predictable and hence derived. A well-known example concerns the relation between syllable-structure and segment-structure information. As observed by Levin (1985), Borowsky (1986) and others, it is often the case that the former can be extrapolated from the latter with a facility equal to that with which the latter can be extrapolated from the former.

Another issue concerns the desirability of having underspecified representations as addresses for lexical storage and retrieval. The despecification of lexical feature values may be viewed as a type of archiving program which compresses information for compact storage.[1] As noted by Lass (1984: 205), Mohanan (1991) and others, one of the motivations for proposing redundancy-free lexical representations in generative phonology seems to have been the assumption that long-term memory constraints prompt speakers to limit storage to idiosyncratic information and to maximise the computing of predictable information. This view has never been seriously defended in the psycholinguistic literature.[2] But if underspecification is not justified by storage considerations, nor does it constitute a particularly plausible model of efficient lexical access. The goal of maximal economy of lexical representation can only be achieved at the expense of greatly increasing the amount of computation to be performed at retrieval (Braine 1974). Just as an archived computer text file has to be de-archived before it can be accessed, so would a speaker-hearer have first to 'unpack' the condensed, underspecified form of a lexical entry before submitting it to articulation or recognition.

---

[1]We owe this analogy to Jonathan Kaye (pc).

[2]This issue is not addressed in recent psycholinguistic applications of feature underspecification (e.g. Lahiri & Marslen-Wilson 1991, Stemberger 1991).

This last point leads on to what is perhaps the most fundamental objection to the notion that the function of phonological derivation is to prepare cognitively represented objects for phonetic implementation. Bromberger & Halle express this notion quite succinctly when they make the claim that systematic phonetic representations 'are only generated when a word figures in an actual utterance' (1988: 53). It is as well to be quite clear about how this view relates to the Chomskyan dichotomy between I-language and E-language. In equating the notion of generation with speech production, it immediately places the phonological component outwith the domain of grammar proper. It is not just that the systematic phonetic level constitutes a buffer between the representation of internalised phonological knowledge and its articulatory or perceptual externalisation; the phonological component as a whole is geared towards fulfilling a goal which is essentially extragrammatical, that of turning out utterance-bound phonetic forms.

The validity of this view cannot be taken for granted. In what follows, we will consider the consequences of adopting an alternative position, in which phonology remains on the competence side of the competence-performance divide. According to this view, phonological processes, in line with general principles, do no more than capture generalisations regulating alternations and distributional regularities. This function can be served quite independently of any provision that needs to be made for articulation and perception. That is, processes can be construed as purely generative in the technical sense of specifying the membership of the set of grammatical phonological structures in a language. Under this view, initial and final representations in phonological derivation are isotypic: processes map phonological objects onto other phonological objects rather than onto phonetic ones. Such a view places phonology firmly in the grammatical camp.

If phonological processes map like onto like, it follows that initial representations should be no less phonetically interpretable than final representations. In fact, initial representations can in principle be envisaged as being wholly indistinguishable in kind from final representations. Such an arrangement is of course impossible if non-final representations are subject to underspecification. But it is perfectly consistent with an approach in which phonological primes enjoy autonomous interpretability. In a full-blown element approach, there is no sense in which a final representation is any more physical or concrete than an initial one. There is thus nothing corresponding to a systematic phonetic level, since any representation at any stage of derivation is directly mappable onto physical phonetics.

It is necessary to bear this point in mind when comparing the empirical content of element theory with that of feature underspecification. It is sometimes assumed that the two approaches somehow share the same realisational outcome, namely a systematic phonetic representation consisting of fully or almost fully

specified matrices of bivalent features (e.g. Pulleyblank, in press). In the light of the foregoing discussion, it is clear that this opinion is mistaken. The assumption gives the incorrect impression that one theory is directly translatable into the other, an impression that has hardly been discouraged by the occasional practice of explicitly spelling out the phonetic exponence of elements in traditional feature terms (Kaye, Lowenstamm & Vergnaud 1985, Rennison 1990). These researchers are quick to point out that features employed in this way serve no more than a phonetic implementation function and as such are not accessible to the phonology. However, it seems to us better to return to Anderson & Jones' (1974) basic proposal that features have no place in element theory whatsoever. Not only does invoking features muddy the water as far as presentation of elemental phonology is concerned; in several respects, it actually hamstrings attempts to provide an accurate account of how elements are mapped onto physical phonetics, a point we will return to below.

2.2 Jakobson *redivivus*

This last point introduces yet another motive for dispensing with references to orthodox features in specifying the phonetic exponence of elements. The use of SPE-type features entails acceptance of the notion that phonological primes are mapped in the first instance onto the articulatory dimension, in spite of what is usually claimed about a generative grammar being neutral between speaker and hearer. This orientation is perhaps most vividly illustrated in the close congruence that exists between current conceptions of feature geometry and Browman & Goldstein's (1989) gestural model of speech production (Clements 1992). To the slight extent that researchers working within this tradition have concerned themselves with the hearer side of the equation, it has been assumed that the primarily articulatory features can be mapped, albeit indirectly, onto acoustic and perceptual dimensions. (For a recent proposal along these lines, formulated within feature geometry, see Clements & Hertz 1991.)

It would perhaps be unfair to speculate that the articulatory slant of much feature theory is not entirely unconnected to the clear articulatory bias of most introductory courses in phonetics. Nevertheless, it is worth pointing out that a belief in the centrality of speech production, however tacit, implies varying degrees of commitment to the motor theory of speech perception, the notion that the listener decodes speech by some species of internal articulatory synthesis. The theory, in various guises, has met with rather less than general agreement in the phonetic literature (see Klatt 1987 for discussion). Many phoneticians and phonologists remain to be convinced of the wisdom of abandoning the Jakobsonian insight that

the phonetic exponence of subsegmental primes should in the first place be defined in acoustic terms.[3] The speech signal, as Jakobson was wont to point out, is after all the communicative experience that is shared by both speaker and hearer (Jakobson, Fant & Halle 1962: 13). Its primacy in phonetic interpretation should hardly be in question, at least if we are to pay more than lip service to the idea that generative grammar is neutral between production and perception.

It is for this reason that the specifications of elements provided in the next section are couched in primarily acoustic terms (an orientation long associated with Dependency Phonology). That is not to say that elements should be construed as acoustic (or articulatory) events. They are properly understood as cognitive objects which perform the grammatical function of coding lexical contrasts.[4] Nevertheless, continuing the essentially Jakobsonian line of thinking, we consider their phonetic implementation as involving in the first instance a mapping onto sound patterns in the acoustic signal. Viewed in these terms, articulation and perception are parasitic on this mapping relation. That is, elements are internally represented pattern templates by reference to which listeners decode auditory input and speakers orchestrate and monitor their articulations.

3 Elements for vowels

3.1 Elemental patterns: [A], [I], [U]

The phonological evidence supporting recognition of the elements [A], [I] and [U] is reasonably well established. For example, we may refer to the pivotal role played by the independent manifestations of these elements in the organisation of phonological systems. The 'corner' vowels *a, i* and *u* figure with much greater frequency than any other segments in the vocalic systems of the world's languages (Maddieson 1984). Moreover, there is a substantial body of distributional and alternation evidence supporting the conclusion that primes of this nature are individually accessible to phonological processing. Thus we find harmony processes, for instance, which address natural classes identifiable with dimensions such as 'palatality' (characterised by presence of [I]), 'labiality' ([U]), or 'height'

[3]There have always been at least some dissenting voices against the Gadarene rush from acoustic to articulatory features, Andersen's being a particularly eloquent example (1974: 42-43).

[4]As Phil Carr has pointed out to us, this means that phonetic events cannot be considered tokens of element types. Assuming such a relationship of instantiation would imply that the two sets of entities were of the same ontological status. This view is incompatible with the notion that elements, unlike physical articulatory and acoustic events, are uniformly cognitive.

([A]). (For examples of all three types of harmony process, see van der Hulst & Smith 1985, van der Hulst 1988.)

Two other vowels are derivable by possible [A]-[I]-[U] combinations of the type assumed in (1): [I,U] (*ü*) and [A,I,U] (*ö*). Further dimensions of vocalic contrast, including those involving nasality and ATR/peripherality, call for two additional elements, one of which we introduce in the next section. Additional distinctions are derived by recognising segment-internal relations of dependency, which capture the notion that the phonetic manifestation of a compound expression reflects the preponderance of one element over others. (For a review of the relevant literature, see Harris & Lindsey, forthcoming.)

How are [A], [I] and [U] to be defined, and how do we derive the results of their combination? For reasons explained in the previous section, we begin by proposing for them definitions which may be mapped rather directly onto the acoustic signal. Before embarking on such an exercise, it is as well to ensure that we disabuse ourselves of a common misconception — the idea that quantitative values specified with reference to waveforms or spectrographic displays are to acoustic phonetics what qualitative categories specified in terms of vocal tract diagrams are to articulatory phonetics. In fact, a more apt proportion is one that relates a spectrogram to a three-dimensional X-ray movie of vocal-tract gymnastics. The closest acoustic analogues of the relatively abstract categories associated with vocal-tract diagrams are idealised spectrographic patterns. (The nearest precedents are the 'pattern-playback' diagrams pioneered by Haskins Laboratories; see, for example, Cooper et al 1952.)

It usually goes without saying that the articulatory specification of phonological representations is appropriately characterised in terms of qualitative categories rather than in terms of the continuously varying quantitative values encountered in speech production. By the same token, the specification of the acoustic signatures of phonological categories should be couched qualitatively in terms of overall quasi-spectral shapes, and not quantitatively. It would therefore be misguided to express the resonance characteristics of elements such as [A], [I] and [U] as quantitative values relating, say, to formant frequencies. Rather it is necessary to determine the gross quasi-acoustic shapes — what we may call elemental **patterns** — which constitute these categories, and for which symbols such as [A], [I] and [U] are no more than shorthand notations (Lindsey & Harris 1990). The grossness of these patterns should presumably be related to their interpretational autonomy, which they exhibit regardless of whether they appear alone or in compounds.

Figure 1. Elemental patterns: [A], [I], [U] (from Harris & Lindsey 1991). The contours are schematic spectral envelopes, plotted here in frames which map onto acoustic spectral slices (vertical axis: amplitude, horizontal axis: frequency). Thus the pattern for [A] specifies energy of lesser amplitude at the bottom and top of the frequency range than in the middle. The precise structuring of the massed energy in the middle of the frequency range is not a criterial part of the definition of this pattern and hence is left blank in the diagram. [U] by contrast is characterised by energy of greater amplitude in the lower part of the frequency range (its precise structuring again not criterial) than in the higher.

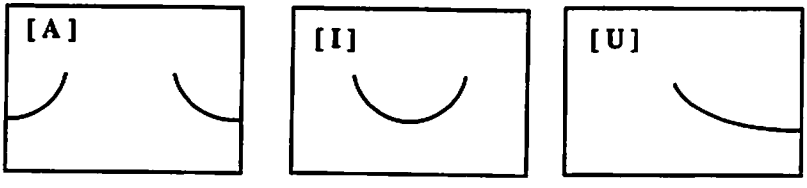
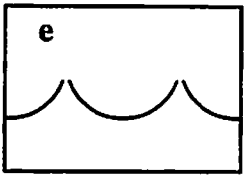
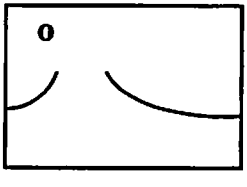


Figure 2. Compounded elemental patterns: (a) [A, I] *e*, (b) [A, U] *o* (from Harris & Lindsey 1991).

(a)



(b)



In Figure 1 we diagram the elemental patterns which, we have proposed elsewhere, characterise [A], [I] and [U] (Harris & Lindsey 1991). We display each pattern in a frame mimicking a spectral slice in which the vertical axis corresponds to intensity and the horizontal axis to frequency. The latter coincides with what we may term the **sonorant frequency zone**, the frequency band containing the most significant information relating to vocalic contrasts (roughly speaking between 0 and 3 kHz).

The elemental pattern of [A], shown in Figure 1a, is appropriately labelled **mAss**. The signal specification of the vowel *a*, the element's independent manifestation, is a mass in the middle of the sonorant frequency zone (ultimately interpretable as the convergence of Formants 1 and 2); that is, energy is higher in the middle of the zone than at top and bottom.

The spectrum of *i* contains a low first formant coupled with a spectral peak at the top of the sonorant frequency zone, the latter peak being relatable to the convergence of Formants 2 and 3. This configuration, with energy lower in the middle of the zone than at either side, may be taken to correspond to a **dIp** elemental pattern characterising [I].

The signal evidence relating to *u* indicates what we may term a **rUmp** elemental pattern for [U]. The vowel displays a peak at the lower end of the sonorant frequency zone (produced by a convergence of the first and second formants); that is, there is no significant energy above the middle of the zone.

The effects of compounding elements are derived by overlaying elemental patterns on one another. The two complex profiles depicted in Figure 2 result from fusing pairs of patterns shown in Figure 1. Figure 2a shows *e*, the outcome of fusion between [A] and [I]. The profile here, which might be described as 'dIp within a mAss', can be viewed as an amalgam of two patterns: (i) energy higher in the middle band than at top and bottom, indicating the presence of [A]; and (ii) energy lower in the middle than on either side, the pattern associated with [I]. In *o*, a compound of [A] and [U] (Figure 2b), we see both (i) energy higher in the middle band than at top and bottom, i.e. [A], and (ii) no energy above the middle, i.e. [U].[5]

Elemental patterns are templates which hearers endeavour to detect in speech input and speakers endeavour to match in the production and self-monitoring of speech output. When a given element is input to speech production mechanisms, the speaker will marshal whatever articulatory resources are necessary or available

[5]Precise modelling of the computations by which the speaker-hearer detects such patterns in acoustic signals is of course no trivial matter. But we assume it is no easier, and probably more difficult, to model detectors of [high], [low], [back] and [round] in X-ray movies of natural speech articulation.

for the spectral realisation of the target elemental pattern. Speech production mechanisms naturally incorporate many preferred specifications for articulatory harnessing. For example, the desired acoustic effect associated with the mAss pattern [A] can be achieved through maximal expansion of the oral tube and constriction of the pharyngeal tube. Note that the imaginary point of maximal height of the tongue body, long cited in impressionistic phonetics (at least since Sweet 1877) as the crucial reference point in the definition of articulatory categories such as 'high' and 'low', has no particular importance in the specification of how [A], or indeed any element, is produced. As is well known, a single vowel sound can be produced with widely varying tongue body contours by different individuals and even by the same individual on different occasions. (See for example Ladefoged et al. (1972) and the results of bite-block experiments of the type reported by Folkins & Zimmerman (1981).) It is by manipulating the overall shape of the vocal airway that the speaker targets particular acoustic effects.

The articulatory incarnation of dIp [I] calls for maximal expansion of the pharyngeal cavity and maximal constriction of the oral cavity. The articulatory implementation of rUmp [U] involves a trade-off between maximal expansion of both the oral and pharyngeal tubes. Labial activity (rounding and/or protrusion, for example) is but one of the factors that contribute to the overall size of the oral cavity. One implication of this is that lip rounding is not an articulatory prerequisite in the production of vocalic expressions containing [U].[6]

Consideration of the articulatory implementation of elemental patterns helps underline the distinction between element-based and current feature-based approaches to the specification of phonetic detail. Feature underspecification refers to the lexical suppression of properties that are phonologically represented at later stages of derivation. As noted earlier, this notion is incompatible with the autonomous interpretation hypothesis, since specified properties are in most instances uninterpretable without support from temporarily suppressed properties. Implicit in element theory, on the other hand, is the conclusion that certain properties assumed by feature theory are nonspecified. The nonspecification of a property implies that it has no representational status whatsoever, either lexically or during derivation. In fact, some such properties, it can be argued, do not even exist as independent entities in physical phonetics. Note that nonspecification, as understood in this sense, in no way impairs the interpretational autonomy of elements. The element [A], for example, has a direct and independent physical interpretation that can be defined without so much as even a passing reference to

[6]This means that our occasional labelling of [U] for convenience as 'labiality' must be taken with an especially large pinch of phonetic salt. Jakobson's term flatness, regrettably out of fashion except among Dependency Phonologists (Anderson & Ewen 1987), is much to be preferred.

features such as [±back], [±low], [±sonorant], [±consonantal], or whatever. Recognition of this point in recent element-based research is reflected in the abandonment of the earlier notion (outlined in Kaye, Lowenstamm & Vergnaud 1985, for example) that phonetic implementation involves the translation of each element into a traditional distinctive feature matrix. Such a featural halfway house, it is now acknowledged, is both logically and empirically redundant.

A further consequence of this view is that there is no place in element theory for anything resembling the featural interpretation of the traditional phonemic notion of contrastivity. A familiar feature matrix includes values whose primary function is to distinguish the sound it specifies from other sounds with which it is in opposition. The matrix for *a*, for example, might include [-round] (whether lexically present or filled in by redundancy rule), which helps differentiate the vowel from a round vowel such as *u*. In element theory, on the other hand, *a* is identified solely on the basis of the only element of which it is composed, namely [A]. To be sure, there is a descriptive sense in which a sound may be viewed as entering into Saussurian relations of contrast with other sounds in a given system; but this does not necessarily imply that each such distinction is directly coded in representation as a particular unit of segmental content. In element terms, the definition of *a* does not include reference to properties that identify other vowels with which it happens to be in contrast, such as the [U] present in *u*. That is, *a* is nonspecified with respect to the pattern characteristics that are relevant to the definition of *u* (acoustically, concentration of energy in the low frequencies, typically realised in articulation by lip rounding). No specification of [U]-related characteristics ('non-rUmp', 'non-round', or whatever) ever enters into the identification of this vowel. Not being *u* is no more a criterial property of *a* than not being an orange is a criterial property of a banana.

3.2 The neutral element

Occurring singly or in combination with one another, [A], [I] and [U] help define the basic set of vocalic contrasts given in (2). Additional elements are necessary for the definition of other dimensions of contrast, including that of peripherality, to which we now turn.

The spectral peaks associated with *a*, *i* and *u* are inherently large and distinct. From an articulatory point of view, this is a reflection of the fact that these vowels, the universally limiting articulations of the vowel triangle, represent extreme departures from a neutral position of the vocal tract. The supralaryngeal vocal tract configuration associated with the neutral position approximates that of a uniform tube and produces a schwa-like auditory effect. The resonating

characteristics of this configuration are such that the first three formants are fairly evenly spaced, with the result that it lacks the distinct spectral peaks found in *a*, *i* and *u*. Most researchers within the element-based tradition accord this neutral quality some special status, either by treating it as a segment devoid of any active elementary content or by taking it to be the manifestation of an independent element, which we will symbolise here as [@]. Broadly speaking, this corresponds to the centrality component in Dependency Phonology (Lass 1984, Anderson & Ewen 1987), to the 'cold' vowel of Government Phonology (Kaye, Lowenstamm & Vergnaud 1985), and to an 'empty' segment lacking any vocalic content in Particle Phonology (Schane 1984) and in the work of van der Hulst (1989).

The element [@] may be thought of as a blank canvas to which the colours represented by [A], [I] and [U] can be applied. From a production point of view, this metaphor reflects the point just made that [A], [I] and [U] are realised by means of articulatory manoeuvres that perturb the vocal tract from its neutral state. From a signal point of view, this implies that the dispersed formant structure of [@] constitutes a base-line on which the elemental patterns associated with [A], [I] and [U] are superimposed.

The idea that [@] is signally blank may come as a surprise to those who assume that the phonologically relevant *dramatis personae* of vocalic acoustics are formants. To be sure, schwa-like vowels exhibit formants just as much as other vowels. However, in the light of what was said in §3.1, it is an assumption we should have no truck with. It is elemental patterns that are the phonologically relevant actors to be detected in vowel signals. Their absence from the dispersed acoustic spectrum associated with [@] indicates a stage on which the phonetic sound and fury of [@]'s formants phonologically signify nothing.

In phonemic and quasi-phonemic transcription, 'ə' is frequently employed as a cover symbol to designate a vowel-reduction reflex. The range of phonetic qualities indicated by the symbol in the literature is quite impressive; in terms of the traditional vowel diagram, it covers varying degrees of openness, backness and roundness. (In transcriptions of Catalan, for example, it symbolises a relatively open value, in Moroccan Arabic relatively close, and in French front rounded.) Some aspects of this variability are phonologically insignificant, but others evidently involve distinct phonological categories. In the former case, variability across languages can be taken to reflect indeterminacies in the fixing of the base line on which other resonance components are superimposed. From a speech production viewpoint, this variability is sometimes characterised in terms of different articulatory or vocal settings. (For a review and discussion of the relevant literature, see Laver (1980).)

In element theory, the independent realisation of [@] may be understood as covering that area of the traditional vowel diagram which is non-palatal, non-open

and non-labial. Non-peripheral categories that are potentially distinct from this base line can then be thought of as displaced versions of the neutral quality. In terms of their segmental make up, these different reflexes can be characterised as compounds in which [@] is fused with some other element(s). The relatively open ə of Catalan, for example, can be represented as a combination of [@] and [A].

The idea that [@] defines the base line on which other element are superimposed can be implemented by assuming that it does not reside on an independent autosegmental tier. Rather it is omnipresent in segmental expressions but fails to manifest itself wherever it is overridden by any another element(s) that may be present. Viewed in these terms, reduction to a centralised vocalic reflex does not involve the random substitution of one set of elements by [@]. Rather it consists in the stripping away of elementary content to reveal a latently present [@].

One of the advantages of viewing centralisation in this way is that it unifies the representation of the process with that of certain processes of raising and lowering which, although not involving reduction to non-peripheral reflexes, nevertheless occur under the same prosodically weak conditions. A widespread phenomenon in the world's languages is a tendency for mid vowels to be banished from prosodically recessive nuclear positions. In metrical systems, recessiveness refers to positions of weak stress; in harmony systems, it refers to nuclei whose harmonic identity is determined by an adjacent dominant nucleus. Under such conditions, it is common to find neutralisation of vocalic contrasts in favour of either non-peripheral reflexes or the 'corner' vowels *a, i, u* or some mixture of both. Height harmony systems such as those of Pasiego Spanish and various central Bantu languages are of the exclusively peripheral type (see Harris & Lindsey (forthcoming) for discussion of the relevant literature). Non-harmonic cases exhibiting similar patterns of neutralisation include Bulgarian and Catalan, in which raising and centralisation co-occur. In Bulgarian, the stressed five-vowel system (*i, e, a, o, u*) contracts to three vowels under weak stress, with the mid vowels raising to high and *a* undergoing centralisation (Petterson & Wood 1987). The stressed seven-term system of Catalan (*i, e*, ɛ, *a*, ɔ, *o, u*) gives way to three terms under weak stress: the *a*-ɐ-*e* contrast is neutralised centrally, the ɔ-*o*-*u* contrast is neutralised under *u*, and *i* remains as it is (Palmada 1991: ch 2).

The fact that the processes just mentioned, raising or lowering of mid vowels and centralisation, all potentially occur in the same general recessive context indicates that we are dealing with a single phenomenon. Although this commonality has long been recognised, it has not always been clear how it should be captured formally. In terms of element structure, however, the processes in question are uniformly expressible as decomposition; all involve the total or partial suppression of segmental material.

Summarising the foregoing discussion of [@], we may identify two main respects in which it differs from other elements. First, it lacks the distinct peak-valley patterns of [A], [I] and [U]. Second, it is latently present in all segmental expressions. One question that remains is how the latter notion is to be accommodated in the mechanism of element fusion. According to Kaye, Lowenstamm & Vergnaud (1985), the desired result can be achieved by assuming that the only circumstances under which [@] contributes anything to the phonetic interpretation of a compound segment are when the element acts as the head of the expression. It will thus fail to make its presence felt in any expression in which it occurs as a dependent. The only circumstances under which latent [@] can become audible are when it is promoted to headship as a result of other elements in the expression undergoing suppression or relegation to dependent status.

4 Summary

Elements are monovalent entities which enjoy stand-alone phonetic interpretability throughout derivation. This conception of phonological primes informs a view of derivation which is in the strict sense generative and which dispenses with a systematic phonetic level of representation.

Elements are cognitive categories by reference to which listeners parse and speakers articulate speech sounds. They are mappable in the first instance not onto articulations but rather onto sound [*sic*] patterns. They constitute universal expectations regarding the structures to be inferred from acoustic signals and to be mimicked in the course of articulatory maturation. No phonological process, we claim, requires modality-specific reference to auditory or vocal anatomy. In particular, we deny any need to recapitulate the latter in the manner of feature-geometric biopsies.

References

Andersen, Henning (1974). Towards a typology of change: bifurcating changes and binary relations. In John M. Anderson & Charles Jones (eds.), *Historical linguistics II: theory and description in phonology, proceedings of the First International Congress on Historical Linguistics, Edinburgh 1973*, 17-60. Amsterdam: North-Holland.

Anderson, John M. & Jacques Durand (eds.) (1987). *Explorations in Dependency Phonology*. Foris: Dordrecht.

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

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

Archangeli, Diana (1984). *Underspecification in Yawelmani phonology and morphology*. PhD dissertation, MIT. Published 1988, New York: Garland.

Archangeli, Diana (1988). Aspects of Underspecification Theory. *Phonology* 5. 183-205.

Archangeli, Diana & Douglas Pulleyblank (1992). Grounded phonology. Ms. Tuscon, AZ: University of Arizona/ Vancouver, BC: University of British Columbia.

Borowsky, Toni J. (1986). Topics in the lexical phonology of English. Unpublished PhD dissertation. Amherst, MA: University of Massachussetts.

Boyce, Suzanne E., Rena A. Krakow & Fredericka Bell-Berti (1991). Phonological underspecification and speech motor organisation. *Phonology* 8. 219-236.

Braine, Martin D.S. (1974). On what might constitute learnable phonology. *Language* 50. 270-299.

Bromberger, Sylvain & Morris Halle (1988). Why phonology is different. *Linguistic Inquiry* 20. 51-70.

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

Clements, George N. (1992). Phonological primes: features or gestures? Ms. Paris: Institut de Phonétique. To appear in *Phonetica*.

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

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

den Dikken, Marcel & Harry van der Hulst. (1988). Segmental hierarchitecture. In van der Hulst & Smith 1988 Part I, 1-78.

Donegan (Miller), P.J. (1978). On the natural phonology of vowels. Unpublished PhD dissertation, Ohio State University.

Jacques Durand & Francis Katamba (eds.) (in press). *New frontiers in phonology*, Harlow, Essex: Longman.

Folkins, J.W. & G.N. Zimmerman (1981). Jaw-muscle activity during speech with the mandible fixed. *Journal of the Acoustical Society of America* 69. 1441-1444.

Goldsmith, John A. (1985). Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology* 2. 251-274.

Harris, John & Geoff Lindsey (1991). Segmental decomposition and the signal. Paper delivered to the London Phonology Seminar, University of London. To appear in *Phonologica 1992: proceedings of the 7th International Phonology Meeting, Krems, 1992*. Cambridge: Cambridge University Press.

Harris, John & Geoff Lindsey (forthcoming). The elements of phonological representation. To appear in Durand & Katamba, in press.

Hulst, Harry van der (1988). The geometry of vocalic features. In van der Hulst & Smith 1988 Part II, 77-125.

Hulst, Harry van der (1989). Atoms of segmental structure: components, gestures and dependency. *Phonology* 6. 253-84.

Hulst, Harry van der (in press). Radical CV phonology. To appear in Durand & Katamba, in press.

Hulst, Harry van der & Norval Smith. (1985). Vowel features and umlaut in Djingili, Nyangumarda and Warlpiri. *Phonology* 2. 275-302.

Hulst, Harry van der & Norval Smith (eds) (1988). *Features, segmental structure and harmony processes, Parts I & II*. Dordrecht: Foris.

Jakobson, Roman, Gunnar M. Fant & Morris Halle (1962). *Preliminaries to speech analysis*. Cambridge, MA: MIT Press.

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

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud (1990). Constituent structure and government in phonology. *Phonology* 7. 193-232.

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

Klatt, Dennis H. (1987). Review of selected models of speech perception. In W. Marslen-Wilson (ed.), *Lexical representation and process*, 169-226. Cambridge, MA: MIT Press.

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

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

Lass, Roger (1984). *Phonology: an introduction to basic concepts*. Cambridge: Cambridge University Press.

Laver, John (1980). *The phonetic description of voice quality*. Cambridge: Cambridge University Press.

Levin, Juliette (1985). A metrical theory of syllabicity. Unpublished PhD dissertation, MIT.

Lindsey, Geoff & John Harris (1990). Phonetic interpretation in generative grammar. *UCL Working Papers in Linguistics* 2. 355-369.

Maddieson, Ian (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Mohanan, K.P. (1991). On the bases of Radical Underspecification Theory. *Natural Language and Linguistic Theory* 9. 285-325.

Palmada Félez, Blanca. (1991). La fonologia del català i els principis actius. Doctoral thesis, Universitat Autònoma de Barcelona.

Petterson, Thore & Sidney Wood (1987). Vowel reduction in Bulgarian and its implications for theories if vowel reduction: a review of the problem. *Folia Linguistica* 21. 261-279.

Pierrehumbert, Janet & Mary Beckman (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.

Pulleyblank, Douglas (1986). *Tone in Lexical Phonology*. Dordrecht: Reidel.

Pulleyblank, Douglas (in press). To appear in Durand & Katamba, in press.

Rennison, John (1984). On tridirectional feature systems for vowels. *Wiener linguistische Gazette*. 33-34, 69-93. Reprinted in Jacques Durand (ed.), *Dependency and non-linear phonology*, 281-303. London: Croom Helm.

Rennison, John (1990). On the elements of phonological representations: the evidence from vowel systems and vowel processes. *Folia Linguistica* 24. 175-244.

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

Stemberger, Joseph Paul (1991). Radical underspecification in language production. *Phonology* 8. 73-112.

Sweet, Henry (1877). *A handbook of phonetics*. Oxford: Henry Frowde.