Variation in Element Theory*

Phillip Backley
Tohoku Gakuin University

This paper introduces the Element Theory approach to segmental structure, and describes the variation that exists between different versions of the theory. Elements are unlike traditional features in several respects: they have only positive values, they represent categories motivated solely by phonological behavior, they map on to acoustic patterns in the speech signal, and they are segment-sized units that can be pronounced independently. The standard version of Element Theory recognizes six elements, giving the phonology a level of expressiveness that is capable of capturing most contrasts, natural classes and phonological processes without overgenerating significantly. Furthermore, standard Element Theory compares favorably with two other forms of the theory, conservative and progressive. These employ different element inventories, which distinguish them from the standard theory at a superficial level. Fundamentally, however, all versions of Element Theory are united by a shared conceptual approach and a common set of assumptions and structural principles.

Keywords: phonology; elements; features; segmental structure; head-dependency; generative restrictiveness; natural classes; consonant-vowel unity; markedness

1. Introduction

This paper describes the main characteristics of Element Theory (hereafter ET), a theory of phonological representation which uses elements, rather than traditional features, to represent the internal structure of segments. It also examines some of the variation that exists between different versions of ET.

At first sight, elements such as |U| and |?| appear to have much in common with features such as [+round] and [−cont], in that both describe segment-internal structure and both are employed as the smallest units in phonological representations. In this sense, ET does not depart significantly from standard feature theories, other than by introducing an alternative vocabulary. However, a closer comparison between elements and features reveals a number of conceptual and
functional differences, which will be outlined in the next section and then further illustrated as the discussion proceeds. These differences reflect the fact that ET began as a response to certain aspects of feature theory that had been considered problematic in one way or another.

It is the collaborative work of Jonathan Kaye, Jean Lowenstamm and Jean-Roger Vergnaud (hereafter KLV) which marks the starting point of ET as an independent research program. In KLV (1985) elements are introduced as “primitives of phonological systems” which function as “autonomous, independently pronounceable units” (KLV 1985: 306). The elements described by KLV are not altogether new, however, since they closely resemble the units of segmental structure (‘components’) used in Dependency Phonology (Anderson & Jones 1974) and the units (‘particles’) of Particle Phonology (Schane 1984). Conceptually, ET even shares some of its core assumptions with more ‘abstract’ theories of segmental structure such as Radical CV Phonology (van der Hulst 1995, 2005). Moreover, many of ET’s developments echo similar developments that have taken place in these related approaches. So it is clear that ET does not exist in isolation. Rather, it is rooted in an established theoretical tradition; historically, this tradition has been most closely identified with work in Dependency Phonology (Anderson & Ewen 1987).

During ET’s relatively short history, its set of elements has been subject to a number of revisions. Some revisions have involved the introduction of new elements; others have led to the merger of two existing elements into a single category, while others have resulted in elements being abandoned altogether. Changes have also been made to the way elements combine to form compounds. These developments will be examined below. Although still relatively small, the ET literature has grown significantly in recent years. For snapshots of ET taken at different points in its development, see Backley (2011), Backley and Nasukawa (2009a), Botma (2004), Botma et al. (2011), Brockhaus (1995a), Charette and Göksel (1996), Cyran (1994, 1996, 2010), Harris (1990, 1994), Harris and Lindsey (1995), Huber and Balogné Bérces (2010), van der Hulst (2000, 2011b), Kaye (1990, 2000, 2001), KLV (1985), Nasukawa (2005), Nasukawa and Backley (2005, 2008), Pöchtrager (2006), Rennison (1998), Ritter (1997, 2005), Sheer (1999a, 2004a, b), Szigetvári (1994), Živanović and Pöchtrager (2010) and references therein.

The paper is organized as follows. Section 2 presents an overview of ET in its standard form, introducing the elements themselves, describing the nature of element representations and outlining ET’s position in relation to phonetics, phonology and the grammar. Then Section 3 looks at the way some versions of ET have departed from the standard model, describing how non-standard versions of ET broadly divide into two types, conservative and progressive. Finally, the discussion in Section 4 concludes that variation in ET is mostly superficial,
involving small differences in the choice of elements or in the way elements combine: beneath the surface all versions of ET share a common set of assumptions, and together, form a unified approach to the representation of segment structure.

2. What are elements?

2.1 Elements and markedness

Elements differ from standard features in several ways. One difference is that elements are unary or single-valued while features are mostly binary. And because elements are unary, ET expresses segmental contrasts by referring to the presence versus the absence of elements, e.g. stops contain |ʔ| (the stop/occlusion element) whereas continuants do not. In feature systems, on the other hand, two segments can contrast by having opposing values of the same feature, e.g. stops are [−cont] while continuants are [+cont]. On this basis, elements may be seen as being inherently positive (but see below) whereas features can be either positive or negative.

An element is positive in the sense that it marks the presence of a given property, never its absence. As a result, it is sometimes impossible for ET to refer to a group of sounds that would traditionally be defined by a negative feature. For instance, while [−nasal] identifies the set of oral sounds, ET has no equivalent: the class of nasals can be represented using the nasal element |L| (cf. [+nasal]), but ET cannot refer explicitly to the complement set of non-nasals. (In ET, oral sounds are oral merely because they lack the positive property of nasality.) This is intentional, however. By restricting its formalism so that it only represents positive properties, ET is able to make fairly strong claims regarding (i) what can constitute a natural class and (ii) what can function as an active or marked property in phonological processes.

In positing the nasal element |L|, for example, ET makes the claim that nasal sounds form a genuine natural class, the members of which all share a common property represented by |L|. This makes it possible for the grammar to refer to nasals as a single class, as illustrated by the numerous phonological processes that specify nasals as triggers or targets. By contrast, the grammar cannot refer to non-nasals as a group because there is no (negative) element to define them. This is consistent with the fact that non-nasals such as [ʃ ʎ ɓ ʁ …] form a random collection of sounds with no shared (positive) properties. And because the grammar cannot represent non-nasals as a class, phonological processes cannot make reference to these sounds exclusively. Again, this is consistent with the facts: non-nasal does not function as an active property in (e.g. assimilation) processes in the way that nasal does. This difference between element-based and
feature-based systems may be restated in terms of markedness. Since elements are positive and identify properties with the potential to be phonologically active, it may be said that the ET grammar deals only with marked properties. By contrast, feature theories deal with both marked and unmarked properties since features can be positive or negative; indeed, standard feature theories make no formal distinction between marked and unmarked. This is not an ideal state of affairs, however, as marked and unmarked properties do not have the same status linguistically.1

From what has just been said, it would seem that ET has an advantage over traditional feature theories in terms of generative restrictiveness, since ET is, by design, limited to describing processes that involve marked properties (e.g. nasal assimilation via active [L]). By contrast, feature theories appear to be looser in that they predict additional processes involving unmarked properties, many of which will never be observed (e.g. oral assimilation via active [−nasal]). Indeed, ET scholars often cite generative capacity as an argument for favoring elements over features. However, there are several points to bear in mind here. First, ET cannot achieve a higher level of generative restrictiveness than feature theories simply by employing unary elements. It must also control the overall number of elements it uses – if the element set is too big, then generative capacity will still exceed desired levels. In fact, this issue is a well-known one, and monitoring (and wherever possible, reducing) the size of the element inventory has always been a priority in the ET program.2

Second, identifying marked and unmarked properties is not always a straightforward matter, as illustrated by the way that some properties which are marked in ET are treated as unmarked in feature-based approaches. For example, ET views open/low as a marked vowel property: all non-high vowels contain |A| because they display some degree of openness. This means that ET is unable to refer to high vowels such as [i u] as a natural class because their only common property is the absence of |A| (and elements cannot have an overt negative value). Moreover, the prediction is that high vowels do not participate actively as a single group in phonological processes – something that appears to be contradicted by feature-based rules that refer to [+high]. Of course, ET has strong supporting reasons for choosing low over high as the marked property, but this choice comes at the expense of having to find alternative ways of analyzing patterns that are traditionally understood to involve the high vowels as a natural class.

A similar situation emerges with the front-back dimension, where disagreement exists over whether front or back is linguistically relevant. ET treats front as the marked property, representing it as the |I| element: front vowels and palatal consonants all contain |I|. From this it follows that backness is not a linguistic property in ET, because it cannot be defined in positive terms. The
Onus is therefore on ET to reanalyze segmental patterns in which [+back] is traditionally viewed as the active property. Furthermore, to determine whether front or back (and similarly, high or low) is marked, it is necessary to observe cross-linguistic patterns of phonological behavior. And unfortunately, data from one language do not always support the same conclusions as data from another language: in some systems front may appear to be active, in others the active property may be back, and in others both front and back could be independently active. In other words, the decision as to whether a property should be marked or unmarked is not always straightforward – it is arbitrary to the extent that the analyst’s choice of supporting data is arbitrary. So even if ET claims to have generative capacity under control, it is necessary to ask whether anything approaching the ‘correct’ set of elements has been used to achieve this level of restrictiveness.

2.2 The phonetics and phonology of elements

Elements differ from features in another respect too: whereas features are based largely on the phonetic properties of speech, elements are to be understood primarily as phonological units that represent the linguistic properties of language. Some element theorists even go as far as to claim that elements are not directly related to actual speech events.

There is a long tradition of using phonological features to describe segment structure, most of the features in current use resembling those found in *The Sound Pattern of English* (Chomsky & Halle 1968), hereafter SPE. And it is clear from the names of these features (e.g. [±high], [±voice], [±lateral], [±anterior], [±round]…) that they are based on phonetic properties – in particular, on properties of articulation such as tongue position, lip position and glottal state. In general, then, features describe segments from the speaker’s point of view. Of course, articulatory features have a linguistic function too, as they identify natural classes and describe phonological processes. But their form suggests that they are ultimately grounded in phonetics, not phonology. The same can be said of features in use before SPE, though with one major difference: these were based on the auditory/acoustic properties of speech rather than on articulation. For example, Jakobson and Halle (1956) employ nine pairs of opposing features (e.g. [compact]/[diffuse], [grave]/[acute], [tense]/[lax]) defined in terms of acoustic properties such as frequency and amplitude. Like SPE’s articulatory features, the auditory/acoustic features proposed by Jakobson and colleagues functioned as linguistic (i.e. contrastive) units. But again, it is clear that they were based on certain phonetic properties of spoken language rather than on phonology.
From the outset, ET has rejected the idea that phonology is directly influenced by phonetics. Instead, it has argued that the units of phonological structure (i.e. elements) should be defined primarily in phonological terms, and that a segment’s representation is determined by studying its phonological behavior, not by listening. This is a surprising claim, given that infants construct lexical (including phonological) forms by listening to the language around them. However, ET’s position begins to make sense once it is understood that the theory distinguishes between a representation, which is a cognitive object consisting of elements, and the realization of that representation, which is an observable event involving physical changes in, for example, air pressure and movements of the vocal organs. The claim is that representations are part of the grammar whereas phonetic knowledge is not. As such, it is possible for someone to acquire their native language even if they lack the ability to speak (e.g. for physiological or psychological reasons). This puts spoken language on a par with writing, to the extent that both provide ways of realizing and communicating linguistic knowledge yet neither constitutes linguistic knowledge itself.

In that case, how does ET see the relationship between elements (as cognitive objects) and spoken language (as a physical phenomenon)? Ultimately, elements must be associated with phonetic properties in some way, otherwise infants would not be able to acquire their native lexicon and mature language users would not be able to use speech as a vehicle for communicating linguistic information. In fact, the elements of ET are associated with acoustic patterns in the speech signal – specifically, with acoustic patterns that are known to be relevant to language (Williams 1998; Harris & Lindsey 2000; Brockhaus & Ingleby 2002; Schwartz 2010). This bias towards the acoustic properties of speech puts ET in touch with the Jakobsonian tradition, which, as mentioned above, favored features defined in auditory/acoustic terms. It also highlights the fact that ET was first developed as a challenge to the widespread use of articulatory features, which proponents of ET considered to be inadequate as a means of explaining segmental behavior.

The reason why elements refer to the speech signal, rather than to some other aspect of spoken language such as articulation, is that the speech signal is the only aspect of the communication process that is shared by speakers and listeners: speakers manipulate their vocal organs in order to create sound patterns in the speech signal, while listeners recover those same patterns from the speech signal and decode them into meaningful language. In this way ET aims to capture the linguistic knowledge common to all users, speakers and listeners alike. The ET literature describes the link between phonology and phonetics as a mapping process whereby the elements contained in a word’s representation are first mapped on to acoustic patterns in the speech signal (Harris & Lindsey 1995). Speakers
then pronounce words by using these patterns as phonetic targets; their task is to reproduce these patterns, and they are free to use their vocal apparatus in whichever way they choose in order to do this. Importantly, the grammar plays no part in this reproduction process. For example, it does not provide the speaker with (articulatory) phonetic instructions on how to achieve these targets – recall that phonetic information in ET is assumed to be external to the grammar. Instead, it is through their experience as talkers (after experimenting with language sounds as infants) that speakers learn to manipulate their speech organs in a way that allows them to reproduce the target acoustic patterns accurately. This scenario takes on board the idea that there is often more than one way of using the vocal organs to reproduce a given acoustic pattern. For example, the pattern for the element |U| is an overall lowering of formant values, which may be achieved through lip rounding, or tongue retraction, or some combination of the two.

In the case of listeners, the decoding of speech proceeds in the expected manner. First, listeners perceive linguistically significant acoustic patterns in the incoming speech stream. They then associate these patterns with the relevant phonological categories, which are represented by elements. And when the resulting element representations are paired with a corresponding lexical representation, word recognition succeeds. Once again, elements function solely as mental constructs while (acoustic) phonetic information relates exclusively to grammatical-external domains such as speech perception. This is true even in the case of derived forms. In an element-based grammar, phonological processes convert (input) mental representations into other (output) mental representations. In principle, both are pronounceable because both can be mapped on to speech signal patterns in the way described above; but in practice, it is only output forms which are pronounced because only these are grammatical (i.e. generated by the grammar). From an ET viewpoint, the role of phonology is to generate well-formed representations, not to convert cognitive objects into something more physical or pronounceable.

An advantage of associating elements with acoustic rather than articulatory properties is that the same elements can be used to represent both consonants and vowels (Backley & Nasukawa 2010). Some degree of consonant-vowel unity had already been achieved by Jakobson and his contemporaries using acoustic features (Jakobson et al. 1952); for example, the feature pair [compact]/[diffuse] distinguishes between open and high vowels and also between back (e.g. velar) and front (e.g. labial, dental) consonants. But following the publication of SPE, when articulatory features began to be used in place of acoustic features, interest in consonant-vowel unity faded. This was because articulatory features are mostly unable to express the properties that are shared by consonants and vowels, forcing SPE’s authors to employ one set of features for consonants (e.g. [±ant], [±cont]) and another set for vowels (e.g. [±tense], [±back]). As a result, properties that are
related on an intuitive level, such as labiality ([+labial]) and rounding ([+round]),
are not related on a formal level. Clearly, this causes some important phonological
generalizations to be missed. It is true that articulation-based feature theories have
since begun to move in the direction of consonant-vowel unity (Clements 1991;
Clements & Hume 1995), but with limited success – although consonant place and
vowel place can be expressed using shared features, most other features remain
exclusive to consonants or vowels because typically the two are articulated in quite
different ways. In acoustic terms, however, the phonetic similarities between con-
onsonants and vowels are more apparent. This will become clear in §2.4, where it will
be shown how the acoustic properties associated with each element can be used to
categorize both consonant categories and vowel categories.

2.3 Element representations

For many, ET is inseparable from the broader theoretical framework known as Gov-
erment Phonology (hereafter GP), a restrictive theory of representations in which
all structural units in a representation enter into asymmetric (head-dependent)
licensing relations. It is through licensing – in some cases, a special form of licensing
called government – that a unit’s presence in a representation is sanctioned. Licens-
ing applies to the units of prosodic structure, such as syllable constituents, and also
to the units of segmental structure; and because the tradition in GP is to use ele-
ments rather than features to represent segments, ET has become closely associated
with GP. Indeed, many see ET as a sub-theory within the GP approach.

In keeping with the nature of the GP framework, ET incorporates head-
dependency relations into segmental representations: an element compound
(i.e. an expression containing more than one element) typically consists of a head
element plus one or more dependent (or operator) elements. In general, heads
make a greater contribution to an overall expression than do dependents, both
acoustically and phonologically. So the most prominent or salient properties of an
expression are usually those belonging to its head element. Compare, for example,
the vowel expressions in (1c) and (1d), which contain the same elements but differ
in headedness.5

(1) a. \([I]\) [i]
   b. \([A]\) [a]
   c. \([I\,A]\) [ɛ]/[ɛ]
   d. \([I\,A]\) [æ]

The mid vowel [ɛ]/[ɛ] in (1c) combines an [I] head with an [A] dependent; this
characterizes it as an [I]-type vowel, but one that has been modified by the addition
of [A]. Its identity is primarily that of a palatal: acoustically it is closer to [i]
(the realization of |I| alone) than to [a] (the realization of |A| alone), and phonologically its most salient properties are those associated with the class of palatal/front vowels. By contrast, [æ] in (1d) is headed by |A| and therefore has the character of an |A|-type or low vowel. Note that the presence of a dependent |I| in |I A| gives [æ] an additional palatal or fronted quality, because speakers must advance the tongue in order to reproduce the acoustic cues (e.g. high F2) associated with |I|.

In expressions containing just one element, this element is usually taken to be headed by default, hence |I| in (1a) and |A| in (1b). This is a reasonable assumption because, if a segment has only one marked property, then that property is presumably its most salient characteristic. On the other hand, some versions of ET depart from this convention by allowing expressions to be non-headed. This approach echoes the practice in Dependency Phonology of recognizing a relation of mutual dependency between co-occurring elements, effectively giving both elements equal status. As can be seen by comparing the contrasts in (2a) with those in (2b), this move increases the number of expressions that can be generated.

(2)  
\begin{align*}
\text{a. } & |U A| \quad [o] \text{ or } [o] \\
& |U A| \quad [\gamma] \text{ or } [o] \\
\text{b. } & |U A| \quad [o] \\
& |U A| \quad [o] \\
& |U A| \quad [\gamma]
\end{align*}

But whether or not ET actually needs non-headed expressions in its vocabulary is a question on which disagreement persists. Some element theorists claim that non-headed structures are necessary for representing certain consonant contrasts, even if they are not required for vowels. Others disallow non-headed expressions altogether, arguing instead for expressions containing a headed neutral element |@|. The status of |@| will be discussed in §3 below.

The preceding paragraphs make it clear that the close association between ET and GP is not only historical but also conceptual, since the notion of head-dependency is central to both. But this does not necessarily mean that ET is inherently bound to GP in the way that is often assumed – indeed, there is no principled reason why element-based representations cannot be employed in other theoretical approaches. That said, however, it seems that the advantages of using elements in place of features can only be fully exploited in frameworks that are based on head-dependency. To explain this point, it should be understood that segmental structure in ET is partly controlled by prosodic structure: strong prosodic positions are able to license strong segments whereas weak positions can only license weak segments. The strength of a position is determined by the head-dependency relations holding between prosodic constituents, constituent heads being strong and dependents weak. For example, branching onsets are conventionally analyzed as left-headed structures, making the left-hand position in (3a) stronger than the right-hand one.
As a relatively strong position, this left-hand slot can support relatively strong segmental expressions, where segmental strength is measured in terms of structural complexity. Broadly speaking, strong segments can be more complex (i.e. they can contain more elements) than weak segments. On this basis it is easy to see why the \( tr \) cluster in (3b) is well-formed whereas \( rt \) in (3c) cannot form a branching onset – in the ill-formed (3c) \( r \) occupies the onset head position but it fails to license its dependent \( t \) because \( t \) (with three elements) is more complex than \( r \) (with just one element). The classic formulation of segmental complexity is given in Harris (1990), who calculates complexity in terms of simple element counting – dependents must be no more complex than their heads. On the other hand, Scheer (2004b) claims that only the resonance elements \( |I|, |U|, |A| \) count towards complexity, since these are (in his model of ET) the only elements shared by all segment types. Another alternative would be to say that, in the case of branching onsets such as \( pl, kr, nj, tw \), it is only the non-resonance elements \( |H|, |L|, |\text{?}| \) that contribute to complexity, since it is precisely these elements that do not appear in the dependent right-hand position. However the details are formalized, the notion of segmental complexity continues to serve as a valuable tool for explaining phonological patterns. Notice, however, that complexity serves as a meaningful concept only because elements are unary and only when they are employed in a phonological framework where representations are based on head-dependency relations.

2.4 The elements

2.4.1 The standard inventory

Although ET exists in various forms, a standard version of the theory has evolved which recognizes the six elements listed in (4). The acoustic and phonological properties shown for each element will be described in more detail below.

(4)  

<table>
<thead>
<tr>
<th>acoustic properties</th>
<th>phonological categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>(</td>
<td>I</td>
</tr>
<tr>
<td>(</td>
<td>U</td>
</tr>
<tr>
<td>(</td>
<td>A</td>
</tr>
<tr>
<td>(</td>
<td>H</td>
</tr>
</tbody>
</table>
Variation in Element Theory

|L| low-frequency energy  | |A| fully voiced obstruents, low tone vowels
|?| sustained drop in amplitude  | |U| oral/nasal/glottal stops, laryngealized vowels

Note that (4) is not a full list of the elements found in ET – some versions of the theory employ additional elements which are not shown here, while other versions use a smaller element inventory than this. Variation in the use of elements will be explored in §3 below.

2.4.2 Resonance elements

The elements |I|, |U|, |A| form a natural group of resonance elements because they all refer to the resonance properties (or formant structure) of speech sounds. They resemble the units of segmental structure proposed in Anderson and Jones (1974), in which |I| represents frontness or palatality, |U| represents backness or rounding and |A| represents lowness or pharyngeality. The same units were later incorporated into Dependency Phonology (van der Hulst 1987, 1989), Particle Phonology (Schane 1984) and ET. As elements, |I|, |A|, |U| first appear in Vergnaud (1982) and then are further developed in KLV (1985). It should be noted that the acoustic descriptions of |I|, |U|, |A| given in (4) are not to be taken as precise or measurable phonetic properties; rather, they are general acoustic patterns which listeners identify in the speech signal and associate with linguistic information. For example, although |I| is identified by a high second formant, the exact frequency of F2 is unimportant – what matters is that it is high enough to merge with F3 and produce a concentration of energy in the F2–F3 region of the spectrum. It is this F2–F3 energy peak which identifies |I|.

Vowels and glides are represented by resonance elements because these sounds are mainly characterized by their formant structure. In addition, |I|, |U|, |A| represent resonance properties (or, expressed in articulatory terms, place properties) in consonants. Vowel and consonant resonance may consist of a single resonance element (simplex resonance) or a combination of elements (compound resonance); either way, headedness can affect the phonological and/or acoustic character of the element(s) concerned. Examples of vowel resonance are given in (1) and (2) above,9 while the main consonant resonance categories are represented as in (5) (adapted from Backley 2011). Again, head elements are underlined.

(5)

<table>
<thead>
<tr>
<th>category</th>
<th>examples</th>
<th>category</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>dental</td>
<td>t, θ, r</td>
<td>A</td>
</tr>
<tr>
<td>I</td>
<td>palatal</td>
<td>tʃ, j, n</td>
<td>I</td>
</tr>
<tr>
<td>U</td>
<td>velar</td>
<td>k, x, ɣ</td>
<td>A</td>
</tr>
<tr>
<td>U</td>
<td>labial</td>
<td>p, Ʉ, m</td>
<td>U</td>
</tr>
<tr>
<td>A</td>
<td>alveolar</td>
<td>ts, s, ʃ</td>
<td>A</td>
</tr>
</tbody>
</table>

All rights reserved
The notion of consonant-vowel unity mentioned in §2.2 is clearly in evidence here, since [I], [U], [A] are equally relevant to consonant and vowel representations. This position is supported by phonological evidence in which consonant place and vowel place properties interact, or otherwise, behave in parallel. In Mapila Malayalam, for instance, a word-final [i] is rounded to [u] after a labial consonant (6b) or a rounded vowel (6c).

(6) a. [dresi] ‘dress’ *[dress], *[dressu]
    b. [jappu] ‘pound’ *[japp], *[jappi]
    c. [onnu] ‘one’ *[onn], *[onn]

In ET terms it is unremarkable that labiality can be reinterpreted as rounding when it occurs in a nucleus (7a), because labials and rounded vowels have the same place specification [U]. But in terms of SPE-type features the same process must be regarded as accidental, as there is nothing to formally connect trigger and target: as (7b) shows, labial consonants are specified as [+ant,−cor] while rounded vowels are [+round].

(7) a. [U] j a p p V
    b. [+ant,−cor] [U] j a p p V

The [I] element is associated with (dental) coronal and also palatal resonance, the difference between the two being expressed by headedness. Headed [I] is present in palatals such as [tʃ jʃ nʃ] and also in complex palatals including the alveolo-palatals [tɕ c z] and the palatovelars [ɕ ɕ z]; in these categories the acoustic pattern for [I] (an F2–F3 energy peak) is especially prominent. In its non-headed form, meanwhile, [I] is the resonance value in some coronal sounds. The presence of [I] in these coronals is backed by phonological patterns in which coronals and front vowels are reported to interact (Hume 1994; Nasukawa & Backley 2011). For example, Scheer (1999a) concludes that [l] and [n] in German have the compound resonance [I A] because both cause [X] to palatalize to [ç] (e.g. [mlç] Milch ‘milk’, [mançɔ] manche ‘many’) in the same way that front vowels do (e.g. [iç] ich ‘I’).

By itself, non-headed [I] is phonetically realized as a coronal tap [r], while a sole headed [I] represents a palatal glide [j] (i.e. palatal resonance with no accompanying consonantal properties). Like some coronals, palatals may also interact with front vowels – a pattern which ET captures by assuming that the [I] element is present in both.

Consonants containing the [U] element display a low spectral peak, which results from a general lowering of all formants. And because lip rounding is an
effective means of bringing about formant lowering, $|U|$ is associated with labial consonants (and also rounded vowels). All bilabials [p b ą β m] have headed $|U|$ as their resonance element. In languages with labiodentals in addition to bilabials, the former is represented by combining headed $|U|$ with a dependent $|A|$, hence $|H A U|$ for [f] versus $|H U|$ for [ɻ].10 By itself, headed $|U|$ is realized as the labial-velar glide [w]. The observation that labial consonants regularly interact with rounded vowels reinforces the assumption that a single property, headed $|U|$, is common to both. In its non-headed form, $|U|$ is phonetically interpreted as velar resonance. Although it may not be immediately obvious that labials and velars are formally related, the historical evidence does point in this direction. For example, sequences of velar+$t$ in Latin were reinterpreted as labial+$t$ in Romanian (e.g. la[k]tem > la[p]te ‘milk’), while in the Skikun dialect of Formosan labials became velars. There is also evidence that labials and velars form a natural class, e.g. in fifteenth century Korean [i] was rounded to [u] before [m p pʰ k kʰ]. Patterns such as these have led standard ET to assume that labials and velars are represented by the same element (Huber 2006; Backley & Nasukawa 2009b). Predictably, a sole non-headed $|U|$ is realized as the velar glide [ʊ]/[ʏ].

Like $|I|$ and $|U|$, the $|A|$ element displays different resonance properties depending on whether it is headed or not. Headed $|A|$ is present in guttural consonants such as pharyngeals [hɨ] and uvulars [q ɢ χ ϵ n]; Bellem (2007) also provides evidence for $|A|$ in emphatics. It will be recalled from $§2.1$ that the $|A|$ element also defines the class of non-high vowels, from which we might expect some non-arbitrary phonological connection to hold between guttural consonants and low vowels. In fact a number of cases are reported in which the two are seen to interact phonologically. In dialects of Arabic, for example, pharyngeal consonants have a lowering effect on neighboring vowels (McCarthy 1994), where lowering presumably involves the addition of $|A|$ to a vowel’s structure. As a non-head, $|A|$ represents (typically, alveolar) coronal place (Ploch 1993; Williams 1998; Kaye 2000; Kula 2005). For example, Kaye (2000) uses the structures in (8a) to represent coronals in Mandarin Chinese, while Kula (2005) represents Bemba coronals as in (8b).

\[ (8) \quad \begin{array}{ll}
\text{a. Mandarin Chinese} & \text{b. Bemba} \\
|t| & |H ʔ A| \\
|n| & |L ʔ A| \\
|s| & |H A| \\
|l| & |ʔ A| \\
|r| & |A|
\end{array} \]

And because coronals contain (non-headed) $|A|$, they too have the potential to bring about lowering on adjacent vowels, e.g. $r$ in Standard Southern Swedish
Lindau 1985), rhotics in Ibero-Romance (Bradley 2011). In fact it is liquids, and especially r-sounds, which provide the initial motivation for positing |A| in coronals. First, most liquids are produced as coronals, and second, they display phonological behavior (e.g. as linking consonants) to indicate that they function as the consonantal (i.e. glide) equivalents of non-high vowels, and should therefore be represented with |A| (Broadbent 1999; Backley 2011: 165ff). And by extension, if liquids have |A| then it may be assumed that |A| also inheres in (at least some) other coronals, as there are many phonological processes which refer to a single class of coronals that includes the liquids.

2.4.3 Non-resonance elements
The remaining elements |H|, |L|, |?| are associated with patterns in the speech signal that refer to properties other than resonance. For this reason they appear mainly in consonants, though in keeping with the spirit of consonant-vowel unity these non-resonance elements can also contribute to the representation of vowels.

The speech signal pattern for the ‘stop’ element |?| involves a sudden and prolonged drop in acoustic energy; this is brought about by a stoppage to the airflow somewhere in the oral cavity or at the glottis. |?| is therefore present in oral stops, affricates and the glottal stop [ʔ]. According to some analyses, |?| also belongs in nasals and laterals because, despite involving continuous airflow, these sounds are produced with an oral occlusion. In fact, the question of whether nasal and lateral stops contain |?| is a phonological one rather than an articulatory one: in some languages they pattern together with obstruent stops as a single class, in which case they presumably have |?| in their structure; but in other languages nasals and laterals pattern with continuants such as fricatives (Mielke 2005), indicating that they do not belong to the class of stops and therefore do not contain |?|.

There are two further points to make regarding the |?| element. First, some versions of ET distinguish between headed |?| and non-headed |?|. Given that heads are acoustically more prominent than dependents, it has been argued that headed |?| identifies the class of glottalic stops comprising ejectives [p’ t’ k’] and implosives [ɓ ɗ Descriptors], in which glottal closure is a prominent feature (Backley & Nasukawa 2009a). Second, in accordance with consonant-vowel unity there are languages that allow |?| to appear in vowels, where vowel expressions with |?| have a laryngealized or creaky quality. For example, Jalapa Mazatec distinguishes between plain (modal voice) vowels [i æ a o u] and laryngealized (creaky voice) vowels [i æ a o u], as illustrated by minimal pairs such as [si] [l] ‘dirty’ versus [si] [ʔ] [l] ‘holiday’ (Silverman et al. 1995). The phonological link between |?| in consonants and |?| in vowels is highlighted in languages such as Capanahua, where creaky vowels alternate with sequences of vowel plus coda |?| (Elias-Ulloa 2009).
The elements |H| and |L| naturally form a sub-group of laryngeal-source elements, mapping on to speech signal patterns featuring, respectively, high- and low-frequency energy. Phonologically, they represent voicing contrasts and nasality in consonants and also tonal contrasts in vowels. In effect, |H| defines the class of obstruents as it is present in (released) stops, fricatives and affricates. In its headed form, |H| is comparable to the feature [+spread glottis] and identifies fortis obstruents in 'aspiration' languages like Swedish, Korean and English. Fortis consonants are produced with a wide glottal opening, and for this reason are typically voiceless; indeed, active voicelessness is an important cue for identifying |H| in consonants. In fortis stops this voicelessness takes the form of voicing lag – a delay in the onset of voicing following stop release, which, in some languages, is interpreted as aspiration. The fortis obstruents [p t k tʃ f θ s ʃ] have |H| in their structure and contrast with their plain or unmarked counterparts [b d ɡ dʒ v ʒ ʒ], which contain no [H] (or [L]). The absence of a laryngeal-source element in plain obstruents renders these sounds phonologically neutral – that is, their laryngeal properties are never active in segmental processes such as assimilation. On its own, headed |H| is pronounced as bare aspiration (or bare frication – see below), i.e. as a glottal [h]. In English, the distribution of [h] mirrors that of stop aspiration (i.e. foot-initial), a pattern that ET captures by representing both as the same structural object |H|. Note also that the glottal fricative [h], unlike oral fricatives, rarely contrasts with its voiced counterpart. The reason why [h] is inherently voiceless derives at least in part from its representation as |H|, this element being associated with fortis (i.e. typically voiceless or aspirated) sounds: in order to reproduce the acoustic pattern associated with |H|, speakers must create a wide glottal opening, which inhibits the vocal fold vibration necessary for voicing.

In its non-headed form, |H| is signaled by the presence of high-frequency aperiodic energy of the kind which characterizes fricatives. In oral fricatives |H| combines with one or more resonance (i.e. place) elements to give structures such as |H A U| [v] and |H A| [s]. ET also assumes that non-headed |H| inheres in released stops – a claim which has both phonetic and phonological support. Acoustically, the release phase of stops is characterized by a period of aperiodic energy similar to that observed in fricatives, albeit much shorter in duration. And although the presence or absence of audible release is never contrastive in stops, the release burst does carry important linguistic information in the form of resonance cues that are crucial for perceiving stops accurately. So in order to reflect the significance of its linguistic role, ET includes |H| in the structure of released stops. Evidence for |H| in stops also comes from consonant lenition. When a stop weakens to a fricative it loses its [ʔ] element, but importantly, any remaining elements may still be realized phonetically. So the elements in the resulting fricative
(i.e. |H| plus resonance) must have already been present in the original stop, e.g. [b] |? H U| may weaken to [β] |(? H U].

The |L| element, which has a function similar to that of [+voice], operates in ‘voicing’ languages like French, Polish and Russian, where the laryngeal distinction in obstruents is between a marked voiced series and a plain (voiceless) series. In (9) the |L|-based (voicing) contrast in French is compared with the |H|-based (voicelessness) contrast in English.

<table>
<thead>
<tr>
<th></th>
<th>VOICED</th>
<th>NEUTRAL</th>
<th>ASPIRATED</th>
</tr>
</thead>
<tbody>
<tr>
<td>(pre-)voiced, unaspirated</td>
<td>[voice] or</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>batte ‘bat’</td>
<td>[bat]</td>
<td>patte ‘paw’ bat</td>
<td>pat</td>
</tr>
<tr>
<td>French</td>
<td>(voicing language)</td>
<td>English</td>
<td>(aspiration language)</td>
</tr>
</tbody>
</table>

In French, obstruents from the voiced series such as [b] in batte ‘bat’ are phonetically fully voiced or even pre-voiced; and because this kind of obstruent voicing has the potential to be active in phonological processes, ET represents it as a marked property using headed |L|. So in voicing languages |L| is found in voiced stops, voiced affricates and voiced fricatives, and these contrast with their plain unvoiced counterparts, which are phonologically neutral because they have no |L| element. Importantly, the split between aspiration languages and voicing languages in (9) is consistent with the known phonetic and phonological similarities that exist between the plain series (transcribed as [b d z ʒ...]) in English and the plain series (transcribed as [p t s ʃ...]) in French. Note that, although |L| functions in a similar way to the traditional feature [+voice], it is not present in sonorants because in these sounds voicing is phonetically spontaneous and phonologically neutral.

In its non-headed form, |L| is identified by an acoustic pattern known as nasal murmur, a broad band of low-frequency energy that can be produced by lowering the velum. |L| is therefore associated with nasals. It inheres in nasal consonants and nasal(ized) vowels, and in addition, functions as an active property in systems of nasal harmony. To phonetically realize |L| on its own, speakers must produce a nasal murmur without superimposing any additional acoustic pattern onto it. This results in a ‘placeless’ nasal of the kind found in Malay and Japanese, which can vary phonetically between a velar [ŋ] and a uvular [ŋ]. Given that
nasality (non-headed |L|) and obstruent voicing (headed |L|) are represented by the same element, it is no surprise that the two show acoustic similarities and also a close phonological relationship. In Japanese, for example, nasal stops behave as though they contain the voicing element |L|, as they trigger voicing assimilation in NC clusters (e.g. *kam-та [kanda] ‘chew (past)’). Yet they also appear to be unspecified for voicing for the purposes of Lyman’s Law, which filters out forms containing more than one segment specified as voiced (e.g. [beni] ‘lipstick,’ [nokogiri] ‘saw’ but *[zabi], *[geda], etc.). Nasukawa (1998) shows how the paradoxical nature of nasals can be accounted for if the same element is taken to be responsible for both nasality and voicing, arguing that it is present as a head |L| in voiced obstruents and as a dependent |L| in nasals. Kula (2002) offers a similar analysis of Bantu voicing dissimilation, where two voiced NC clusters (i.e. nasal plus voiced stop) are prevented from co-occurring in a word. The dissimilation effect itself causes a voiced stop to be reinterpreted as a nasal – a change that may be seen as minimal in that it involves only a switch in headedness from |? Л| (voiced oral stop) to |? L| (nasal stop), e.g. *N-genda [ηηenda] ‘I go’ in Luganda.

Like other elements, |H| and |L| contribute to consonant-vowel unity by appearing in vowels as well as consonants, |H| being phonetically realized as high tone in vowels and |L| as low tone. Contour tones are represented by various combinations of the two, e.g. |H L|, |H L| and |H L| are all potentially contrastive. Support for this dual interpretation of |H| and |L| comes from the known phonetic and phonological link between laryngeal properties and tone (Halle & Stevens 1971). The relevant articulatory dimension is vocal fold (VF) tension, which correlates with fundamental frequency (F0) as shown in (10).

| VF tension | |L| |
|---|---|---|
|VF tension|stiff VFs|slack VFs|
|F0|raised F0|lowered F0|
|in vowels|high pitch|low pitch|
|in consonants (headed)|voiceless|voiced|
|in consonants (non-headed)|frication|nasal|

Because high pitch and voicelessness show similar values for VF tension and F0, ET assumes that each is a different phonetic interpretation of the same structural object, namely |H|. For the same reason, low pitch and obstruent voicing are thought to have |L| in common. The phonetic facts are backed by evidence from phonological patterns in which vowel tone and consonant laryngeal properties are seen to interact in systematic ways (Tang 2008; Yip 2002).
Owing to space limitations, many details concerning non-resonance elements have not been included here. However, (11) summarizes the main consonant categories that can be represented by combinations of $|H|$, $|L|$, $|\text{?}|$ (adapted from Backley 2011).

\begin{center}
\begin{tabular}{|l|l|l|l|}
\hline
\textit{category} & \textit{segments} & \textit{languages} \\
\hline
$|\text{?}|$ & neutral stops & $[p/b, t/d, t/\text{?}]$ & all \\
$|\text{?} H|$ & aspirated stops & $[k^{\text{h}}, t^{\text{h}}, p^{\text{h}}]$ & $|H|$ languages \\
$|\text{?} L|$ & voiced/pre-nasalized stops & $[b, \text{\textit{\textipa{\text{?}}}}d, g]$ & $|L|$ languages \\
$|\text{?} H L|$ & ejectives & $[p', t'f', q']$ & Montana Salish \\
$|\text{?} L|$ & implosives & $[b, d, \text{\textipa{\text{?}}}f]$ & Kimantuumbi \\
$|\text{?} H L|$ & breathy-voiced stops & $[b^{\text{h}}, d^{\text{h}}, g^{\text{h}}]$ & Hindi \\
$|H|$ & neutral fricatives & $[f/y, s/z, s/\text{\textipa{\text{?}}}z]$ & most \\
$|H|$ & voiceless/aspirated fricatives & $[f, f, x]$ & $|H|$ languages \\
$|H L|$ & voiced fricatives & $[v, \text{\textipa{\text{?}}}d, z]$ & $|L|$ languages \\
$|L (\text{?})|$ & nasals & $[m, n, \text{\textipa{\text{?}}}n]$ & most \\
$|L H (\text{?})|$ & voiceless nasals & $[m, \text{\textipa{\text{?}}}n, \text{\textipa{\text{?}}}n]$ & Jalapa Mazatec \\
$|L H (\text{?})|$ & breathy-voiced nasals & $[m, \text{\textipa{\text{?}}}n]$ & Newar \\
$|L \text{?}|$ & laryngealized/creaky nasals & $[m, \text{\textipa{\text{?}}}n]$ & Kwakw’ala \\
\hline
\end{tabular}
\end{center}

Clearly, no single language will make use of all these combinations of $|H|$, $|L|$, $|\text{?}|$ – in reality, most employ only around five or six of the categories listed. This does not mean that ET overgenerates, however, because it is still the case that each category is an observed natural class. What (11) does show is that $|H|$, $|L|$ and $|\text{?}|$, in conjunction with headedness, have enough expressive power to represent a wide range of (even very large) consonant systems.

3. Variation in ET

3.1 Introduction

So far the focus has been on the use of elements in standard ET – a version of the theory that has evolved over time and, consequently, looks rather different from the original ET model of the 1980s. Other versions of ET also exist, however. These alternatives to the standard theory typically follow one of two directions: ‘conservative’ ET (or Con.ET) has retained many of the features of the original model and employs a larger element set than the one described above, while ‘progressive’ ET (or Prog.ET) has departed radically from both Con.ET and standard ET by replacing a number of the elements with structural devices such as inter-element licensing relations.
A paper such as this cannot possibly cover all aspects of variation in ET. However, it can give a flavor of these alternative approaches for the purposes of comparison. Con.ET and Prog.ET are both insightful approaches to segmental structure, yet from the discussion below it will become clear that standard ET as described in §2 manages to strike a useful balance between the two, providing a restrictive model of phonological knowledge in which elements are abstract enough to function as cognitive units of linguistic structure yet concrete enough to be realized phonetically without the need for explicit rules of phonetic interpretation. Furthermore, the set of six elements used in standard ET has the expressive power to describe most segmental contrasts without overgenerating excessively. In order to understand something of the nature of Con.ET and Prog.ET, this section discusses several aspects of segmental phonology that receive different analyses in standard and non-standard versions of ET.

3.2 Nasality

In standard ET nasality and obstruent voicing are represented using the same element – recall from §2.4.3 that non-headed [L] refers to nasality and headed [L] to voicing. As already described, this move helps account for the interplay observed in a number of languages between nasals and voiced stops. For example, Botma (2004) reports how nasality functions as a word-level property in Tukuya, dividing the lexicon into separate sets of oral and nasal words. The voiced stops [b d g] and the nasals [m n ŋ] are thus in complementary distribution: voiced stops appear only in oral words (e.g. [bipi] ‘swollen’, [síge] ‘follow’) and nasals only in nasal words (e.g. [mípí] ‘badger’, [nítí] ‘coal’). By contrast, neutral stops such as [p] and [t] may appear in either set, as they have no [L] element and therefore cannot participate in the oral/nasal division. From the Tukuya facts it can be inferred that voiced stops are the oral equivalent of nasal stops; and if both are assumed to share the same [L] element then they can be distinguished purely by headedness: logically, an element cannot be headed and non-headed at the same time, so nasals (with non-headed [L]) are prevented from appearing in words from the oral set (characterized by headed [L]), and vice versa.

In Con.ET, on the other hand, nasality is marked by a separate nasality element [N], leaving [L] to represent obstruent voicing and low tone. This, in effect, denies the existence of any phonological relation between nasals and voiced obstruents. In one sense this approach is understandable, as the evidence for a link between nasality and voicing (from processes such as voiced obstruent nasalization, for instance) comes from only a minority of languages. But on the other hand it misses a useful generalization that helps to explain why, for example, pre-nasalization targets voiced but not voiceless stops. It would seem that
Con.ET employs a separate nasal element |N| because this was standard practice in the work of KLV. And in turn, KLV assumed the existence of |N| because, although their early ET model was radically different from the SPE feature system it was designed to replace, their thinking was still at least partly rooted in traditional (phonetically-based) segmental classifications, one such classification being the category ‘nasal’. In 1980s phonology, there was no question that nasality had the status of an independent phonological category and that nasal existed as a universally marked property that should be represented by a unique element.

Although the idea of using both |L| and |N| has now been rejected by many working in standard ET, there is perhaps one piece of evidence in its favor which concerns nasal vowels. If nasality is represented by |L|, as it is in standard ET, then nasal vowels must contain |L|, e.g. |A L| [a ˜]. But recall that |L| also represents low tone on vowels. A question therefore arises as to the structure of a nasal vowel which has low tone. To my knowledge, no satisfactory solution to this issue has been proposed in the ET literature. It is conceivable that nasal |L| and low tone |L| could be specified at different levels of structure, one operating as a segmental property and the other having a wider (e.g. word-level) distribution. However, this is something that requires further investigation if the single-element approach to nasality and low tone is to gain general acceptance.

3.3 Labials and velars

In a situation that parallels the one just described, standard ET represents labial resonance and velar resonance using the same element, whereas Con.ET employs two separate elements. Recall from (5) that in standard ET headed |U| marks labials and non-headed |U| marks velars. This shared-element approach captures the acoustic similarities between the two – both are associated with patterns characterized by a low spectral peak – and also the phonological link that appears to hold between them (§2.4.2). Yet in Con.ET this connection between labials and velars is effectively denied by assigning each category its own element: labials have |U| while velars are represented by the ‘neutral’ element |@|.

|@| belongs to the set of resonance elements because its spectral pattern refers to formant frequency – specifically, to non-converging formants of the kind found in central vowels such as [ə]. The |@| element has much in common with the ‘cold vowel’ introduced in KLV (1985), in that neither has any inherent active properties; moreover, both appear to have their origins in the centrality component |ə| of Dependency Phonology (Lass 1984). The |@| element is phonologically neutral in the sense that it displays none of the marked properties associated with the positive resonance elements |I|, |U|, |A|. And as a result, it is not usually phonologically
active in the way that other elements are. Exceptionally, however, [@] does become phonologically significant when it is the head of an expression, because this automatically imposes dependent status on any co-occurring elements.17 And given that heads make a greater contribution than dependents to a compound, it follows that segments with headed [@] should be characterized by the defining property of this element, i.e. centrality (or the absence of the positive resonance properties associated with [I], [U], [A]). The role of [@] in vowels will be discussed in §3.4. The focus here is on consonantal [@], which is phonetically realized in Con.ET as velar resonance.

Harris and Lindsey (1995) give acoustic and phonological arguments for representing velar resonance as [@]. Whereas other resonance categories are associated with specific and stable formant patterns, velars have an acoustic profile that is largely dependent on that of the following sound, suggesting that they do not have a defining resonance property of their own. And for this reason, the neutral element [@] seems appropriate. Phonologically too, there is evidence that velars lack a marked place specification. In ET, consonant weakening is analyzed as the loss of elements from a consonant’s structure (Harris 1990, 1994, 1997; Harris & Kaye 1990; Harris & Lindsey 1995; Harris & Urua 2001); and once an element is lost, it cannot be phonetically realized. But as already noted, speakers may still pronounce any remaining elements in the usual way. The patterns of weakening in (12) show how oral stops are reinterpreted as glides following the loss of their non-resonance elements [?] and [H].

After [?] and [H] are lost, the remaining resonance element is realized as a glide. But in velars the outcome is often zero rather than the expected velar glide [u], e.g. Turkish ine[k]~ine[ ] ‘cow’ (Harris & Lindsey 1995:67). What this indicates, according to Harris and Lindsey and other Con.ET theorists, is that velars have no overt resonance element, unlike other consonants, so that when their non-resonance properties are lost, only empty structure is left.

The arguments for representing velars with [@] are appealing. However, this approach does raise a number of issues. First, if velars are less complex than other resonance categories (because they lack a positive resonance element) then, according to the principles of ET, they must be less marked – indeed, they ought to behave as unmarked sounds. Yet it is coronal that typically functions as default resonance in neutralizing contexts and positions where an epenthetic consonant
is required. Second, if velars have empty resonance there is the question of how the glottal consonants [h] and [?] should be represented, since ET already assumes that glottals lack an active resonance property, i.e. [h] as [H] and [?] as [?]. Moreover, [h] and [?] may function as epenthetic consonants whereas velars typically do not. And third, if labials and velars have different resonance elements then again a useful generalization concerning the phonological link between these categories is missed altogether (see §2.4.2). By contrast, standard ET benefits in the ways outlined above because it recognizes labials and velars as a natural class of [U] segments.

3.4 ATR

The ET literature describes several ways of representing the ATR contrast in vowels. KLV (1985) propose that an independent ATR element [I] is present in ATR vowels but absent from non-ATR vowels, e.g. ATR [I A] [e] versus non-ATR [I A] [ɛ]. So for them it is [I] which defines the natural class of ATR vowels, and in addition, identifies ATR domains in languages with ATR harmony. For a time, versions of Con.ET followed KLV in employing [I], which resembles the ATR component [a] of Dependency Phonology (Anderson & Ewen 1987:305). However, serious doubts about the status of [I] soon began to emerge (van der Hulst 1990; Charette & Kaye 1993), quickly leading to its wholesale abandonment.

KLV (1985) note that the ATR element [I] has unique properties that prevent it from being viewed as a regular member of the set of resonance elements. And in fact, it was this special status of [I] that brought about its early demise. They describe the marked property of [I] as something akin to the feature [+ATR], associating it with pharyngeal resonance: all vowels produced with tongue advancement are therefore assumed to contain [I], as this has the effect of enlarging the pharyngeal cavity. According to KLV, [I] is unique in several respects: unlike other elements it does not normally function as the head of a compound, and also, it is not usually pronounced as a simplex expression. Exceptionally, however, single [I] can be phonetically realized as [i] in languages such as Kpokolo (KLV 1985), i.e. as a vowel which is tense (owing to the presence of [I]), back (absence of [l]), high (absence of [A]) and unrounded (absence of [U]).

When analyzing ATR harmony languages like Kpokolo, it seems reasonable to posit an independent ATR element. But on the other hand, the vowel system of Kpokolo is atypical, because in the majority of the world’s languages ATR is neither harmonically active nor even lexically contrastive. For example, in the unmarked 5-vowel system [i u e o a], ATR plays no active role. Despite this, in the KLV model the phonetically ATR vowels [i u o] must be represented by compound expressions consisting of [I] plus the relevant combination of [I], [U], [A]. This approach has at least two drawbacks. First, it creates an anomalous situation in which languages
with \[i\ u\ e\ o\ a\] (or indeed, just \[i\ u\ a\]) require an ATR element in their grammar, even though ATR is phonologically redundant in these systems. Second, it wrongly implies that a vowel system containing only non-ATR vowels (e.g. \[I\ u\ e\ o\ a\] or \[I\ o\ a\]) is less marked than the more familiar \[i\ u\ e\ o\ a\] system, since ET advocates that an increase in segmental complexity (via the introduction of \(|I|\)) entails an increase in markedness. KLV do address this latter point (KLV 1985: 312ff), but as Roca (1994: 129) notes, their arguments are not especially convincing.18

An alternative way of representing ATR is developed in Harris and Lindsey (1995), following Charette and Kaye (1993), who argue that ATR vowels are headed by \(|I|\), \(|U|\) or \(|A|\) while non-ATR vowels are headed by the neutral element \(|@|\) introduced in §3.3. So for example, ATR \(|I\ A|\ [ɛ]\) contrasts with non-ATR \(|@\ I\ A|\ [ɛ]\). This approach reflects the idea that ATR vowels have cues which place them at the periphery of the acoustic vowel space (i.e. close to the acoustic targets for \(|I|\), \(|U|\), \(|A|\)) whereas non-ATR vowels tend to display less extreme resonance cues, drawing them closer to the central region of the vowel space – recall that centrality and its associated acoustic properties are defining characteristics of \(|@|\).19 Harris and Lindsey assume that \(|@|\) is latently present in all vowels, providing a ‘baseline resonance’ which becomes audible only when it is not masked by the presence of a specified \(|I|\), \(|U|\) or \(|A|\) element. So in effect, \(|@|\) marks the absence of any lexically specified resonance, e.g. \([u]\) may be represented as \(|U|\) or as \(|U\ @|\), where \(|@|\) in the expression \(|U\ @|\) marks the absence of \(|I|\) or \(|A|\). Non-headed \(|@|\) is thus phonologically neutral, functioning as a place-filler: it has no marked properties, and when combined with other resonance elements it makes no contribution to the resulting expression. Structures such as \(|A\ I|\) and \(|A\ I\ @|\) are therefore phonologically and acoustically equivalent. It is precisely for this reason that some ET scholars deny the existence of \(|@|\) as a true element, treating it instead as a marker of empty structure.

By contrast, headed \(|@|\) is phonologically significant because it imposes its salient property, centrality, on expressions in which it occurs. This allows ATR contrasts to be expressed by a difference in headedness alone. For instance, in non-ATR \(|I\ A|\ [ɛ]\) both elements have dependent status – specifically, they are dependents of a headed \(|@|\). Thus \(|I\ A|\) may be rewritten as \(|I\ A\ @|\), and in languages with an ATR contrast this non-ATR expression will be distinct from ATR \(|I\ A|\ [ɛ]\). In a typical (e.g. West African) ATR harmony language, where harmonizing vowels divide into separate harmonic sets, vowels may be represented as in (13).

(13) ATR (headed) non-ATR (non-headed)

\[
\begin{array}{ll}
[i] & |I| \\
[u] & |U| \\
[e] & |I\ A| \\
[o] & |U\ A|
\end{array}
\quad
\begin{array}{ll}
[i] & |I| \text{ or } |I\ @| \\
[o] & |U| \text{ or } |U\ @| \\
[e] & |I\ A| \text{ or } |I\ A\ @| \\
[o] & |U\ A| \text{ or } |U\ A\ @|
\end{array}
\]
Harris and Lindsey then capture ATR harmony in terms of headedness agreement: in ATR domains all harmonizing vowels must be headed by |I|, |U| or |A| (depending on the language), and in non-ATR domains they must all be headed by |@|. Note that to express alternations between ATR and non-ATR, the grammar must allow for the headed/non-headed status of an element to be switched, e.g. if the triggering vowel is ATR (headed) then all harmonizing vowels within the domain must become headed if they are not already so. This is illustrated in (14) by the two forms of the instrumental-locative suffix le in Vata (Mahanta 2007), where the suffix vowel is interpreted as headed [e] |I A| following the ATR stem [pi] ‘prepare’ in (14a) but as non-headed [e] |I A @| after the non-ATR stem [ɓl] ‘sing in’ in (14b).

(14)  
a. pi-le [pi-le] ‘prepare with’ 
b. ɓl-le [ɓl-le] ‘sing in’

The mechanism which ensures that headedness agreement takes place is called head alignment (Harris & Lindsey 1995:65) because, following the traditions of autosegmental phonology, individual elements are assumed to occupy separate tiers and, in the case of ATR domains, all the elements on a given tier must be aligned for headedness. This is illustrated by the structure in (15a).

(15)  
a. [pi-le] ‘prepare with’  
b. [ɓl-le] ‘sing in’

What Harris and Lindsey (1995) demonstrate is that, by employing the neutral element |@|, non-ATR domains can be represented in the same way as ATR domains since both involve head alignment along a given tier: in ATR domains the harmonizing tier contains a specified element while in non-ATR domains the tier is phonologically empty and therefore associated with the neutral (place-filling) element |@|. Thus it is possible to describe ATR harmony even in the absence of an independent ATR element. Note, however, that this approach has the drawback of setting ATR harmony apart from other types of vowel harmony (e.g. those involving active labial/|U| or palatal/|I|), which are uniformly viewed as spreading processes (cf. KLV 1985), whose analysis of ATR harmony in Kpokolo uses a spreading ATR element |f|).
Although standard ET does not recognize \( |@| \) as an element, it is still used in some versions of Con.ET because it clearly adds to the expressive power of the model. Not only does it offer a way of representing ATR harmony, but it also succeeds in capturing the seemingly neutral acoustic properties of velar resonance (§3.3). In addition, it provides a convenient way of referring to the neutral/inactive status of central vowels such as \( [\text{a}] \), and offers a solution to the problem of how to distinguish between the back unrounded vowels \( [u \gamma \lambda] \) (headed by \( |@| \)) and the back rounded vowels \( [u o n] \) (headed by \( |U| \) or \( |A| \)), as noted by Scheer (1999a). Nevertheless, in several respects \( |@| \) appears out of place among the regular elements of ET, and consequently, has been dropped from standard versions of the theory. Unlike the other resonance elements, \( |@| \) is not usually phonologically active (e.g. in assimilation processes), making it difficult to justify \( |@| \) as a marked property in segmental structure. In fact, if \( |@| \) is latently present in all segments, as Con.ET assumes, then it cannot function as a true lexical property in the way that genuine elements do. Also, the phonetic properties associated with \( |@| \) are mostly expressed in negative terms, e.g. ‘central’ resonance tends to be regarded as the absence of other positive resonance properties, rather than as a positive property in itself. Moreover, the central resonance pattern consists of non-converging, equally-spaced formants – quite unlike the patterns associated with other resonance elements, in which linguistic information is equated with energy peaks created by converging formants. All in all, an element inventory without \( |@| \) seems to be more in line with the principles of the ET approach as a whole. Furthermore, the functions of \( |@| \) in Con.ET can be reanalyzed satisfactorily in standard ET by referring either to other elements (e.g. \( |U| \) in velars) or to headedness (e.g. to represent ATR distinctions).

3.5 One ‘H’ or two?

Con.ET departs from standard ET in its use of two separate elements, \( |H| \) and \( |h| \), to represent the laryngeal-source properties of voiceless obstruents. It will be recalled from §2.4.3 that standard ET employs just a single element \( |H| \) to cover (what it considers to be) the related properties of frication, voicelessness and aspiration.

Following the tradition established in early ET, Con.ET recognizes a ‘noise’ element \( |h| \) to represent frication in fricatives and audible release in stops, and also the \( |H| \) element to represent voicelessness in aspiration languages such as English. This produces structures such as those in (16) for English obstruents.

\[
\begin{array}{lcl}
|f| & |U| h & \text{fortis fricative} \\
|p| & |U|^h |H| & \text{released fortis stop} \\
|p^h| & |U|^h |H|^21 & \text{aspirated stop} \\
|p|^h & |U|^h (h) H & \text{unreleased stop}
\end{array}
\]

\( |\text{ish}, \text{bee}|f\) \\
\( |\text{olite, o}|p|\text{en} \)

\( |\text{ace, a}|p|\text{ear} \)

\( |\text{sou}|p, \text{ha}|p|\text{less} \)
As independent elements, |h| and |H| have distinct physical properties: |h| is associated with high-frequency aperiodic energy while |H| is given, somewhat unexpectedly, the articulatory label ‘stiff vocal cords’ (KLV 1990: 216). From the viewpoint of standard ET, separating frication and voicelessness in this way is reminiscent of articulation-based approaches in which even the phonological characteristics of segments are defined phonetically by referring to voice-manner-place categories.

Once again, using two elements instead of one has the obvious advantage of greater expressive power. But at the same time, it is well known that a larger number of elements increases the possibility of overgeneration. In such a situation, theorists must consider the question of cost: does a boost in expressive power outweigh the cost (in terms of potential overgeneration) of having an additional element? A ‘yes’ response would indicate that the extra element is necessary, perhaps because the theory has no alternative means of reaching the desired level of expressive power. On the other hand, a ‘no’ response would suggest that the linguistic information expressed by the extra element can also be expressed using structure or elements already available. In the case of |h|, standard ET has assumed the latter, allowing the merger of |h| and |H| into a single category |H|. The initial impetus for this merger was the significant degree of overlap, both acoustic and phonological, between the elements |h| and |H|.

In contrast to the merger just described, Pöchtrager (2006) argues for a more radical way of streamlining the element inventory and avoiding the overlap between |h| and |H|. Rather than expressing |h| as an element, he proposes that the linguistic information carried by |h| is better expressed as a structural property. For this reason, the version of ET developed in Pöchtrager (2006) may be viewed as a progressive form of ET, or Prog.ET. Conventionally, a single (i.e. non-geminate) consonant occupies just one position, where ‘position’ refers to a terminal node in the syllabic structure. For Pöchtrager, however, two positions are needed to represent fricatives, a head position and its dependent. In fortis fricatives a segmental expression is realized in both positions, while in lenis fricatives segmental material is confined to the head.

\[ (17) \quad \begin{alignat}{2}
\text{a. } & \text{ [v], e.g. } & \text{give} \\
\text{b. } & \text{ [f], e.g. } & \text{laugh} \\
X_1 & \xrightarrow{O'} xO\{U\} \\
X_1 & \xrightarrow{O'} xO\{U\}
\end{alignat} \]

The dotted arrows in (17b) mark a relationship of m(elodic)-command holding between head and dependent (Pöchtrager 2006: 68), where the head position
xO m-commands its dependent x₁. By being m-commanded, x₁ can acquire and phonetically interpret the element expression in its head – in the case of (17b), the [U] element. By contrast, the structure in (17a) involves no m-command relation, leaving x₁ structurally and also phonetically empty. M-command is therefore capable of capturing the fortis/lenis difference in obstruents, which in standard ET is represented by the merged [h]/[H] category described above. In addition, structural complexity (i.e. two positions, as opposed to one) takes the place of [h]/[H] in representing frication, thereby identifying the class of fricatives. Pöchtrager’s initial motivation for replacing [h]/[H] with structural complexity came from his understanding of the way vowel length interacts with voicelessness and frication in English. The structure-based approach of Pöchtrager is examined further in Schwartz (2010), who considers its implications for the phonetic realization of elements, phonological markedness and consonant strength.

Meanwhile, other scholars have also explored ways of expressing non-resonance properties without explicitly referring to non-resonance elements; these include Jensen (1994), Golston and van der Hulst (1999), Ritter (1997), Cyran (2010) and Szigetvári (2004). To take one example, in Golston and van der Hulst (1999) ‘manner’ properties (e.g. continuancy) are encoded in syllable structure rather than segmental structure, while laryngeal properties such as voicing are expressed as modifiers of the syllabic constituents onset, nucleus and coda. Golston and van der Hulst do not dispense with a frication feature altogether, however. Instead, they propose that the resonance properties of a fricative are licensed by a sub-syllabic feature marked F (for fricative). There are also terminal features labeled P (for plosive), R (for sonorant), A (for low) and I (for high). To illustrate their model, (18) shows the structure of the English word *plank* [plæŋk] (Golston & van der Hulst 1999:155).

(18)
In other approaches, Ritter (1997) represents laryngeal-source properties entirely in terms of headedness, claiming that stops and laterals have headed structures, as do nasals (which are headed by |L|), whereas fricatives are non-headed. Cyran (2010: 34ff) develops a similar idea for fricatives in Irish. But rather than rejecting |h| entirely, he argues that its appearance is parametrically controlled, and that Irish is a language in which |h| has no active role. So in order to represent friction in Irish, Cyran too refers to the headedness of |l|, |u| and |A|. Meanwhile, Szigetvári (2004) entertains the possibility of abandoning the ‘manner’ elements |h| and |?| altogether and expressing the same information via governing (licensing) relations between skeletal positions – so, as in Golston and van der Hulst (1999), Ritter (1997) and Cyran (2010), segmental (melodic) information is recast as structural (prosodic) information. In Szigetvári’s approach, for instance, the property traditionally represented by |?| is taken to be the realization of an ungov erned C position. Szigetvári also suggests that the obstruency associated with |h| may be better represented by the presence of a licensed C position. A significant advantage of this approach is that it helps to explain why neither |?| nor |h| is regularly active in assimilation processes: if they are represented as licensing relations rather than as elements, then they cannot spread.

Another version of Prog.ET is developed in Rennison and Neubarth (2003), who propose yet another means of eliminating |h| from the set of elements, leaving |H| to represent aspiration in consonants and high tone in vowels. Again, their aim is to show that the properties of |h| are structural rather than segmental. They do this by making radical use of the neutral element |@|, which they dub the functional element F. Rennison and Neubarth argue that the special status of F allows it to replace the elements |A| |?| |F| |h| of standard ET and/or Con.ET. The F element is not associated with a specific acoustic pattern; instead, it “maximizes or distracts from the most discernible acoustic pattern of the respective syll position to which it is attached. In the case of consonants, maximization means silence (stops) and distraction means noise (spirants/fricatives)” (Rennison & Neubarth 2003: 8–9). The various ways in which F can contribute to element expressions are summarized in (19).

(19)  

<table>
<thead>
<tr>
<th></th>
<th>headed</th>
<th>non-headed</th>
</tr>
</thead>
<tbody>
<tr>
<td>F in consonants</td>
<td>stop (cf.</td>
<td>spirant/fricative (cf.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F in vowels</td>
<td>non-high (cf.</td>
<td>A</td>
</tr>
</tbody>
</table>

What makes their approach interesting is the fact that not only |h| but also |?| is expressed as a structural property, the suggestion being that both |h| and |?| are to be viewed as unmarked properties (in consonants) and not as marked segmental properties. This view is reinforced by the fact that |h| and |?| are not regularly active.
in spreading processes in the way that |I|, |U|, |A| and |L| are known to be: whereas place assimilation and nasal harmony are widespread, for instance, |h| assimilation and |ʔ| harmony are rarely (if ever) observed. Other proposals to redefine |ʔ| as a structural property have been made in Bachmaier et al. (2004), Golston and van der Hulst (1999), Jensen (1994), Kula and Marten (1998), Pöchtrager (2006), Ritter (1997) and Szigetvári (2004). Unsurprisingly, versions of Prog.ET that have replaced |h| and |ʔ| with structural information are left with drastically depleted element sets; for instance, Pöchtrager (2006) works with an inventory of just four elements |I|, |U|, |A|, |L|.

3.6 Vowel resonance

To represent vowels, most element theorists follow standard ET in using the three resonance elements |I|, |U|, |A| together with their headedness properties. However, Scheer (1999a) proposes an additional element |B| to represent rounding (and also labiality in consonants), leaving |U| to represent backness. The relevant segmental categories are shown in (20).23

<table>
<thead>
<tr>
<th></th>
<th>vowels</th>
<th>consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>back</td>
<td>velar</td>
</tr>
<tr>
<td></td>
<td>round</td>
<td>labial</td>
</tr>
</tbody>
</table>

The following structures, adapted from Scheer (1999a:210), illustrate the role of |B| in vowel representations.

(21) a. [U] [u]
    [B] [u]
    b. [U B] [u, o]
    [A B] [o]
    [U A B] [o, o]
    c. [I] [i, i]
    [I B] [y, y] (NB: *|I U|)

In systems with both |U| and |B|, |U| produces backness without rounding to give [u], while |B| produces rounding without backness to give [u], as shown in (21a). So to represent a vowel that is back and round, both elements must be present, hence the structure [U B] for [u] shown in (21b). Employing |B| in addition to |U| affects representations in other ways too; for instance, front rounded vowels must contain |B| rather than |U|, giving the structure [I B] for [y, y] in (21c) (cf. |I U| in standard ET). Moreover, it follows that the elements |I| and |U| may not combine in the same vowel expression because they represent conflicting properties, |I| encoding frontness and |U| backness. Scheer (1999a:211) argues that this is a desirable result because |I| and |U| do not
co-occur in consonants, and if a filter such as *|I U| applies to consonants then ideally it should apply to vowels too.24

The option of distinguishing between |U| (back, velar) and |B| (round, labial) has not been taken up in standard ET; instead, |U| explicitly encodes rounding and, implicitly at least, encodes backness too. This position reflects a general reluctance on the part of ET scholars to recognize new elements of any kind – as already noted, expanding the element set increases the risk of overgeneration while the goal for many working within ET is to develop a representation system in which generative capacity is carefully controlled. Furthermore, an advantage of using a single back/round element |U| is that it highlights the connection between velar/back and labial/round (see §3.3): in articulatory terms there is a link between rounding and backness in vowels and between labiality and a raised (back) tongue position in consonants, while in acoustic terms the two properties clearly overlap (cf. Jakobson’s feature [grave]). By contrast, in versions of ET that employ both |B| and |U|, this link must be viewed as accidental – there is no reason why two independent (and apparently unrelated) categories such as |B| and |U| should display a close connection of this kind.

With regard to vowel resonance, there is another way in which some progressive versions of ET depart from standard ET. Although |I|, |U|, |A| are widely accepted as the basic units of vowel structure, attempts have been made to eliminate |A| from the grammar on the grounds that it displays some unique properties that set it apart from |I| and |U|. For example, |A| is inherently more syllabic or vowel-like than the other vowel elements, causing it to be drawn naturally to the head position of a nucleus. For the same reason, |A| tends to be absent from nuclear dependent positions. Assuming that the head position in a long nucleus is on the left (KLV 1990: 198ff), this accounts for the fact that diphthongs such as [ai] [au] are typologically more common than [ia] [ua]. Notice, furthermore, that [ia]/[ua] are readily reinterpreted as the glide-vowel sequences [ja]/[wa], which indicates how the less vowel-like |I| and |U| (though not |A|) can be realized outside the nucleus as onset glides.

To capture the syllabic nature of |A|, Rennison (1998) redefines |A| in structural terms: if a vowel expression has no element as its head (or in Rennison’s terms, if it is headed by the empty element) then it is realized as though it contains |A|, i.e. as a non-high vowel with a converging F1–F2 acoustic pattern. In this way, the properties usually ascribed to |A| are (i) directly associated with the syllable nucleus and (ii) given an unmarked status because they correspond to the absence of any specified vowel elements. A reanalysis of |A| in terms of structure has also been developed by Markus Pöchtrager and colleagues (Živanović & Pöchtrager 2010), who represent the properties of |A| by means of a branching structure in the nucleus. Although the idiosyncrasies of |A| are widely known, and the motives
for separating |A| from |I| and |U| are well understood, standard ET continues to make regular use of |A| in vowel representations. Evidently, |A| assumes such an established position in the element inventory that any calls for its removal tend to meet with strong resistance. Moreover, the idea that an empty element (for Ren- nison) or an unspecified branching nucleus (for Živanović & Pöchtrager) can be pronounced in such a specific and consistent manner (e.g. as [a]) is a challenging one for most ET practitioners.

3.7 Combining elements

When elements combine, they form an asymmetric relation in which one element functions as the head of the compound while any remaining elements are dependents. And as already noted, heads generally display greater prominence or ‘visibility’ (van der Hulst 2011a) than dependents, both acoustically and phonologically. The obvious advantage of asymmetric fusion is that it increases the number of contrasts that can be expressed using a given number of elements. But in another respect, this advantage can also be a disadvantage because it can lead to overgeneration, multiplying the number of possible expressions beyond what is needed to represent natural languages. The effect of headedness on generative capacity is illustrated below: in (22a) four expressions are generated from just two elements, while in (22b) twelve expressions are produced from three elements.

(22) a. |I| i |A| e |I| i |U| u |U| y
    |A| I| e |A| I| e |A| U| o |I| U A| æ
    |I| A| æ |I| A| æ |U| A| o |I| A U| ø
    |A| a |A| a |I| U| u |U| A I| ø

And although the point is often overlooked, the overgeneration problem is even more serious in the case of consonants, which involve more elements and therefore more head-dependent configurations.

To tackle overgeneration, ET has introduced a number of mechanisms for controlling the way elements combine. KLV (1985) adopt the practice established in autosegmental phonology of having different elements occupy different tiers, where each tier exists as an autonomous plane in a three-dimensional space. And while an element on one tier is free to combine with elements on other tiers, element fusion is restricted by assuming that some elements share the same tier and are thus prevented from combining. For example, KLV give the five-vowel system [i u a e o] as in (23), where ‘v’ is equivalent to |@| in that it marks the absence of a specified resonance element.
To generate this vowel system and at the same time exclude non-occurring vowels, KLV have |I| and |U| share a tier. This means that both elements can combine with |A| to create the mid vowels |I A| [e] and |U A| [o] but cannot fuse with each other. If |I| and |U| were allowed to combine, the result would be either |I U| [y] or |I U| [u], both of which are ill-formed in languages with a symmetrical five-vowel system. Recently, little attention has been paid to tier structure in standard ET; for many ET theorists, tier structure has now been superseded by licensing constraints (see below).

KLV (1985, 1990) also refer to an additional mechanism for controlling element combination called charm. They claim that each element has an inherent charm value – either positive, negative or neutral\(^{25}\) – where charm value broadly equates with relative sonority, and that (positively or negatively) charmed elements of like value are unable to combine. So for example, a vowel cannot be marked for both high tone (H\(^{-}\)) and low tone (L\(^{-}\)) because it is impossible for the two negatively charmed elements H\(^{-}\) and L\(^{-}\) to fuse. And in conjunction with the theory of government (licensing) developed in KLV (1990), charm also accounts for phonotactic patterning. Briefly, the overall charm value of a segment is usually determined by the charm value of its head element, and charmed segments may act as governors whereas charmless (i.e. neutrally charmed) segments occupy governed positions. In addition, KLV stipulate that segments with negative charm cannot occur in a nucleus while those with positive charm are banned from non-nuclear positions. Ultimately the use of charm was short-lived; it was abandoned in all versions of ET by the mid 1990s after it had become clear that, although charm values could describe segmental patterns, they could not explain why those patterns are the way they are.\(^{26}\) Elements had been assigned their charm values on the basis of how they were seen to behave, i.e. in terms of their syllabification and their ability to co-occur with other elements. But as Scheer (1999b) points out, observation does not in itself lead to explanation.

Most versions of ET abandoned the use of autosegmental tiers too, as they also failed to provide an adequate solution to the issue of overgeneration. For example, the tier structure in (23) for the five-vowel system [i u a e o] overgenerates because, as well as producing the grammatical expressions |I A| [e] and |U A| [o], it also generates unwanted compounds such as *|I A| [æ] and *|U A| [v].
which differ from [e] and [o] only in terms of headedness. Note, however, that numerous languages do employ the expressions |I A| [æ] and |U A| [ɒ], which suggests the need for a more precise way of expressing element combinations on a language-specific basis. In fact this need was met through the introduction of licensing constraints, which function as parametrically variable statements controlling (a) the headed/non-headed status of individual elements and (b) licensing relations between co-occurring elements. Examples of both types of licensing constraints are shown in (24).

(24) a. |I| must be a head
   All expressions are headed
   |I| and |U| must be heads
b. Only |h| can license |L|
   |A| is not a licensor
   |? h L| must license operators

Initially, licensing constraints were introduced to ensure that ET would generate the correct segment inventories for a given language (i.e. all the contrastive segments without any non-occurring segments), the idea being that elements could combine freely but only within the bounds of what the licensing constraints deemed grammatical. Later, however, it was found that the same licensing constraints could also account for the outcome of certain phonological processes. For example, Charette and Göksel (1996) show that vowel harmony processes in Turkish take the form they do because of the particular licensing constraints that operate on the (static) vowel system of the language. The use of licensing constraints does raise some questions, however. From the examples in (24) it can be inferred that the form of licensing constraints is very flexible, and consequently, that there must be a large number of possible licensing constraints that could operate in languages but are never actually observed. There is also the matter of acquisition. Because licensing constraints are generalizations over an entire vowel or consonant system, it follows that they cannot be fully acquired until the whole set of segmental contrasts is in place, at which point the language learner no longer has any need for a mechanism for distinguishing native from non-native segments. On this basis it is not altogether clear what role licensing constraints play in the acquisition process. Despite issues such as these, however, licensing constraints continue to be employed in standard versions of ET as a means of controlling overgeneration.

A rather different approach to controlling element combination is taken in van der Hulst (2011b), who does not make use of licensing constraints. Van der Hulst shows how the six elements in his model naturally divide into three pairs, where the members of each pair stand in a polar relation. Vowels and consonants
have a common internal structure comprising three articulatory gestures, which are organized as in (25). Each gesture is then associated with a different pair of elements.

\[(25)\]

\[
\begin{array}{c}
\text{Laryngeal (tone)} \\
\{L, H\}
\end{array} \quad \begin{array}{c}
\text{Manner (aperture)} \\
\{A, \forall\}
\end{array} \quad \begin{array}{c}
\text{Place (color)} \\
\{U, I\}
\end{array}
\]

In each pair, the element written on the left is vocalic (or ‘rhyme-oriented’) in character while the element on the right is consonantal (or ‘onset-oriented’); the vocalic elements \(|L|, \mid A\mid, \mid U\mid\) can therefore be rewritten as \(|V|\) while the consonantal elements \(|H|, \mid \forall\mid, \mid I\mid\) can be replaced by \(|C|\).\(^31\) Note that collapsing three elements into a single category like this does not cause them to merge phonologically because they belong to different gestures and are effectively in complementary distribution.

In van der Hulst’s model, elements combine to form head-dependent relations of the kind used in standard ET. This is true whether the elements in question belong to different gestures or to the same gesture. The latter case is exemplified by contrasts such as \(\mid A \forall\mid \mid x\mid\) versus \(\mid A \forall\mid \mid x\mid\). Dependency relations also hold between gestures, but are universally fixed to the extent that the manner gesture is always the head of the whole expression. In this respect van der Hulst’s approach is more restrictive than standard ET, where any element is a potential head, and consequently, many more possible contrasts based on headedness can be generated. For example, Harris (1994: 126) distinguishes between strident and non-strident fricatives using the headedness of \(\mid h\mid\), e.g. \(\mid h U\mid \mid f\mid\) versus \(\mid h U\mid \mid x\mid\). No such headedness distinction is possible using the structure in (25), because any expression equivalent to \(\mid h U\mid\) (where manner is dependent on place) is ill-formed. For details of how element expressions are phonetically interpreted in van der Hulst’s approach, see van der Hulst (2011b).

Segment-internal organization of the kind shown in (25) appears in other versions of Prog.ET too, particularly those that have been influenced by work in Dependency Phonology. Botma (2004) proposes the structure in (26), where place elements are again dependent on manner elements (i.e. manner dominates place on the vertical axis).
Notice that the elements |H|, |L|, |?| appear twice, once in the manner gesture and again in the phonation gesture. Notice also that manner and phonation differ in terms of prosodic status. Because the manner gesture is dominated directly by a timing slot ‘x’, the manner elements |H|, |L|, |?| function as the head of the segment; as such, they form an obligatory part of the structure. By contrast, the phonation gesture is a dependent of the syllabic constituent (onset or nucleus), and as a dependent, plays more of a secondary role. As phonation elements, |H|, |L|, |?| represent a variety of laryngeal modifications including aspiration and breathy voice (from |H|), voicing and nasalization (from |L|) and glottalization and ejection (from |?|). Besides Botma (2004) and van der Hulst (2011b), other models of Prog.ET have also exploited dependency relations between elements in one way or another – see, for example, Nasukawa and Backley (2005), Kula (2005) and Huber and Balogné Bérces (2010).

4. Discussion

Although ET exists in several versions, the differences between these versions are mostly superficial, typically involving minor variations in the number of elements used and in the way elements are allowed to combine. Beneath the surface, the various forms of ET are fairly similar, being united by a common set of assumptions which, in turn, are motivated by the same grammatical principles and theoretical reasoning. These shared assumptions create a core theory that defines the ET approach as a whole. And it is this core theory that is outlined in §2 and given the label ‘standard ET’ because it is recognized by the majority of ET users. What makes this version of ET ‘standard’ is, above all, its use of six elements. The names of these elements sometimes vary; for example, the name of the combined nasality/voicing element varies between |N| (Nasukawa 2005) and |L| (Ploch 1999), while the name of the combined frication/voicelessness element varies between |H| (Kaye 2000) and |h| (Huber & Balogné Bérces 2010). But importantly, the number of elements remains constant. Within ET the consensus seems to be that six is an ideal number of elements for building representations that provide a workable level of expressive power without excessive overgeneration.
In conjunction with headedness, the six elements are capable of expressing most of the lexical contrasts, natural classes and phonological processes that a segmental theory is expected to handle. And as a result of merging some of the categories that had been assumed in earlier versions of ET, standard ET also manages to capture some important phonological and acoustic associations such as the nasality-voicing relation (via a merged |N|/|L|) and the voiceless-obstruent relation (via a merged |h|/|H|). Another benefit of the six-element model is that it makes consonant-vowel unity a realistic goal, every element having the potential to appear in consonant and vowel expressions. This applies even to the non-resonance elements |H|, |L| and |?|, which, although naturally suited to consonants, represent contrastive vowel properties such as nasality (|L|), tone (|H|, |L|) and creaky voice (|?|). A perceived drawback of standard ET is its apparent phonetic vagueness, e.g. a single phonological expression such as |I A| may be pronounced [ɛ] in some languages but [æ] in others. It should be clear from §2, however, that differences such as this are not strictly a matter for phonology: during acquisition language learners set precise phonetic targets for element expressions, but knowing how to pronounce a given expression forms part of their communicative ability, not their phonological knowledge.

Versions of ET that depart from the standard model differ in the number of elements they use and also in the way elements combine. Broadly speaking, two opposing trends have emerged, referred to here as conservative ET (Con.ET) and progressive ET (Prog.ET). Con.ET retains the character of earlier forms of ET and typically employs a larger number of elements than is found in standard ET.34 This is because Con.ET has mostly resisted attempts to reduce the size of the element inventory. As a result, typical features of Con.ET include

i. separate elements for nasality and voicing (e.g. |N| and |L|)
ii. separate elements for frication and voicelessness (e.g. |h| and |H|)
iii. the use of a neutral element such as |@| for velar resonance and non-ATR vowels.

With ‘extra’ elements such as |N|, |h| and |@|, Con.ET has no difficulty expressing a broad range of natural classes and phonological processes without relying too much on abstract notions such as headedness. For example, by treating nasality (|N|) and low tone (|L|) as separate properties it becomes possible to represent vowel systems where tone and nasality are both contrastive – and importantly, where they function independently of each other. Similarly, if the grammar distinguishes voicelessness (|H|) from noise (|h|) then it can express the difference between voiceless released stops (with |h|) and voiceless unreleased stops (without
Variation in Element Theory

Harris (1994: 124) shows how some laryngeal neutralization patterns (e.g. in Thai and Korean) refer precisely to this difference. And unlike standard ET, Con. ET can distinguish between strident (headed \[|h|\]) and non-strident (non-headed \[|h|\]) sounds independently of any accompanying laryngeal (\[|H|\) or \[|L|\) contrast (Harris 1994: 126).

But on the other hand, there are some relations holding between elements that Con.ET is unable to express easily. For example, voicelessness and aspiration (both represented by \[|H|\) are closely linked to obstruents (i.e. \[|h|\) sounds), yet this connection appears to be coincidental in Con.ET, where \[|H|\ and \[|h|\ are viewed as independent properties. So although a larger element set might lead to greater expressive power, it can have disadvantages: some natural classes are difficult to formalize and some useful phonological generalizations can be missed. Most importantly, however, the possibility of overgeneration increases, leaving Con.ET theorists to rely on formal mechanisms such as licensing constraints to control the grammar’s output.

Prog.ET takes a more radical approach to the element reduction program. As a result, it typically employs fewer elements than standard ET, having replaced some of them with structural properties (e.g. head-dependency) or other representational devices. And by employing representations such as those in (27), it is clear that some versions of Prog.ET assume a greater degree of abstractness than is found in other forms of ET.

(27)  a. \[|p|\ (Pöchtrager 2006)  b. \[|p|\ (Schwartz 2010)

The motivation for replacing elements with structural properties like these is well understood – for some time ET has questioned why some elements appear to be more element-like than others. And indeed, the response from Prog.ET is a reasonable one, in which those elements that display untypical behavior are redefined in non-element terms. But for element theorists working in the Con.ET or standard ET traditions, it can be difficult to see how configurations consisting mainly of organizing structure (e.g. branching constituents, licensing relations) can be mapped consistently on to specific acoustic patterns in the speech signal in the same way that elements can.
No doubt Prog.ET has other interesting developments and challenging ideas in store, designed to streamline the form of representations still further. Nevertheless, the expectation is that standard ET will continue to be viewed as the most accessible and workable approach to element-based structure, offering an explanatory theory of segmental representation that achieves a degree of success in modeling the relation between internalized language and spoken language.

Notes

* My thanks go to Bert Botma, Jeroen van Craenenbroeck and three anonymous reviewers for their valuable comments on an earlier draft.

1. This point had been made by Trubetskoy, Jakobson and other Prague School linguists. Feature theorists have since proposed mechanisms that allow feature systems to incorporate markedness into representations – see Chomsky and Halle (1968) for Markedness Theory and Archangeli (1988) for Underspecification Theory.

2. Early attempts at reducing the size of the element set include Backley (1993), Jensen (1994) and Charette and Kaye (1993). More recent developments are described in §3 below.

3. As Harris and Urua (2001) note, interest in the auditory-acoustic dimension of speech has also been revived in some feature theories, e.g. Flemming (1995) and Boersma (1998).

4. For descriptions of GP and analyses carried out under the GP umbrella, see KLV (1985, 1990), Charette (1990), Harris (1990), Kaye (1990), Brockhaus (1995b), Scheer (1999a), Ritter (2006), Cyran (2010) and references therein. Like ET, GP exists in several versions: recent developments have led to the emergence of distinct structure-based models including the Coda Mirror approach (Dienes & Szigetvári 1999; Ségréal & Scheer 2001, 2008; Scheer 2004a; Scheer & Ziková 2010), VC Phonology (Szigetvári 1999) and Head-driven Phonology (van der Hulst & Ritter 1999).

5. In this paper, head elements are underlined. Note that some versions of ET follow the convention of showing head elements on the right, in which case \( |A.I| \) would be equivalent to \( |I.A| \).

6. I am grateful to Bert Botma (p.c.) for making this point. For a survey of the theoretical approaches that employ head-dependent asymmetries, see van der Hulst (2011a).

7. For an alternative view of constituent licensing, see Scheer (2004a).

8. Harris (1997), following Harris and Kaye (1990), shows how consonant weakening may be expressed as a reduction in complexity via the loss of elements from an expression. See also Harris and Urua (2001).

9. The acoustic properties of vowels are described in Harris and Lindsey (2000).

10. The motivation for assuming an additional \( |A| \) in labiodentals is discussed in Scheer (1999a).
11. For example, Backley and Nasukawa (2009a) propose that headed $|H|$ is common to aspirates and fortis fricatives – the sounds that trigger English sonorant devoicing, e.g. $[k^{(h)}\text{rai}]$ cry, $[f\text{rais}]$ fry.

12. In aspiration languages like English the fortis/lenis distinction in fricatives is captured by the headedness of $|H|$; (voiceless) fortis fricatives have headed $|H|$ whereas (plain) lenis fricatives have non-headed $|H|$ (where $|H|$ is realized as frication, not voicelessness). By contrast, voicing languages like French do not employ headed $|H|$ because voicelessness is not active in these systems; instead, the difference between (voiced) lenis and (plain) fortis fricatives is marked by the presence/absence of the voicing element $|L|$, hence $|H\ L\ A|$ [z] versus $|H\ A|$ [s] (see below for a discussion of the functions of $|L|$).

13. By contrast, unreleased stops lack $|H|$, e.g. $[p\, ð\, U]$ in chapter. In general, the absence of $|H|$ makes it more difficult for listeners to perceive the resonance cues for these stops.

14. In the GP tradition, phonological processes cannot arbitrarily add new elements to a structure. This rules out the possibility of deriving $|H|$ phonologically during the weakening process.

15. The typological split between $|H|$/aspiration languages and $|L|$/voicing languages is discussed in, for example, Harris (1994:133ff), Honeybone (2005), Iverson and Salmons (2006) and Cyran (2010, 2011).


17. This idea was first proposed in KLV (1985) in relation to the cold vowel $|v|$.

18. An anonymous reviewer has pointed out that the question of complexity/markedness in relation to ATR is no longer seen as an issue in ET. This is because element theorists now agree that the theory should allow for some degree of phonetic variation among languages without an ATR contrast, hence $|I|$ as $[i]$~$[i]$, $|I\ A|$ as $[e]$~$[ɛ]$, etc.

19. Note that, when $|@|$ is phonetically realized on its own, it has a value somewhere in the $[a]$~$[i]$~$[u]$ range, depending on the language. Variation of this kind is expected, precisely because $|@|$ corresponds to the absence of $|I|$, $|U|$, $|A|$ and therefore to the absence of any specific acoustic cues.

20. An anonymous reviewer has pointed out that this analysis introduces confusion over the status of $|@|$. If $|@|$ has no positive properties then it should be regarded as a ‘place-filler’ rather than as a true element. Yet as a place-filler there is no reason for $|@|$ to have its own autosegmental tier in the way suggested by (15b). (In ATR harmony languages the resonance elements are conventionally arranged on just two tiers, one occupied by $|A|$ and the other by $|I|$ and $|U|$, from which it follows that the headed tier in (15b) must belong exclusively to $|@|$.)

21. The possibility of allowing expressions to have more than one head is discussed in Backley (2011:196ff).

22. See, for example, Backley and Nasukawa (2009a) and Huber and Balogné Bérces (2010).

23. The difference between $|B|$ and $|U|$ in Scheer (1999a) echoes an earlier distinction made in some versions of Dependency Phonology between the components $|u|$ and $|o|$ (Lass 1984:278ff).
24. Recall from (5), however, that standard ET does permit \(|I|\) and \(|U|\) to combine in consonants, and moreover, represents front rounded vowels such as \([\text{y} \, \partial \, \text{o} \, \text{œ}]\) as compounds of \(|I|\) and \(|U|\) without needing to refer to a separate rounding element \(|B|\).

25. Harris (1990) gives the charm values of elements as follows: positive \((N^{+}, A^{+}, I^{+})\), negative \((H^{-}, L^{-})\), neutral \((I^{0}, U^{0}, R^{0}, \nu^{0}, \theta^{0}, h^{0})\).

26. For a critical review of the notion of charm, see Coleman (1990).


29. The question of elements as polar categories is explored in van der Hulst (2000:92ff). See also Backley (2011:194ff) for a discussion of polar (‘antagonistic’) relations between elements.

30. The \(|\forall|\) element is similar to the neutral element \(|@|\) used in standard ET and the centrality component used in Dependency Phonology, although it has a wider range of phonetic interpretations depending on whether it is (i) nuclear or non-nuclear and (ii) a head or a dependent.

31. In other sources, van der Hulst employs only the elements \(|C|\) and \(|V|\) in representations. See, for example, van der Hulst (2005) for a description of the RcvP (Radical CV Phonology) approach.

32. In Botma (2004) the \(|L|\) element represents sonorancy as well as voicing and nasalization, so it is present in vowels and sonorant consonants. Accordingly, all segments are assumed to contain at least one manner element.


34. See, for example, Cyran and Nilsson (1998) and Harris (1994).

35. See, for example, Pöchtrager (2006), Schwartz (2010) and Živanović and Pöchtrager (2010).

References


All rights reserved


Ploch, Stefan. 1993. The vocalic systems of standard German and Augsburg Swabian and rhotic effects on the internal structure of vowels. MA dissertation, SOAS, University of London.


Variation in Element Theory


Author's address

Phillip Backley
Tohoku Gakuin University
1-3-1 Tsuchitoi
Aoba-ku
Sendai
980-8511, Japan

phillipbackley@tsc.tohoku-gakuin.ac.jp