Consonant-Vowel Unity in Element Theory*

Phillip Backley        Kuniya Nasukawa  
Tohoku Gakuin University          Tohoku Gakuin University

ABSTRACT. This paper motivates the Element Theory view that vowels and consonants are described using a single, unified set of features (elements). This position contrasts with traditional feature-based models in which vowels and consonants are described in their own terms using separate feature sets. These models miss the fact that vowels and consonants can belong to the same natural class and thus share features. We highlight this using data showing a phonological link between tonal properties in vowels and laryngeal properties in consonants. Elements (but not articulatory features) can express a close acoustic connection between the two, because elements make reference to the acoustic properties of speech.

Keywords: Element Theory, acoustic cues, lexical tone, laryngeal contrasts, feature systems

1. Introduction

A basic division between consonants (C) and vowels (V) is assumed throughout the phonology literature. Moreover, theories of segmental structure which follow the traditional feature-based approach tend to adhere to Assumption 1:

Assumption 1
Consonants and vowels should be described in separate terms, since they have obvious articulatory differences and independent phonological characteristics.

The articulatory differences between Cs and Vs have been described countless times, and need not be repeated here. We instead focus on their phonological differences. It is usually the case that a phonological process will target either Cs or Vs, but not both. And for a long time this encouraged feature theories to represent Cs and Vs separately, using different sets of features. Moreover, it was felt that C and V features should exist independently of each other, so most feature systems made it awkward to describe C-V interaction in a natural way.

Unfortunately, Assumption 1 does not always fit the facts, as we readily see Cs and Vs interacting systematically. In response, some theories have tried to abandon Assumption 1 in favour of a style of representation which formally integrates C and V properties. Taken to its natural conclusion, this integrated approach might be expressed as Assumption 2:

Assumption 2
Consonants and vowels are able to interact phonologically because they have descriptive properties (i.e. features) in common.

The problem with Assumption 2 is that it is not supported by the facts of articulation: Cs and Vs are indeed articulated in quite different ways. Therefore, if feature theories are determined to continue describing language sounds in terms of their articulatory properties, then these theories face the real challenge of identifying features which apply simultaneously to both consonant and vowel articulations.
One alternative is to assume that features are not based on articulation at all. This is the view taken in Element Theory (Nasukawa and Backley (2008)). The ‘elements’ (features) of Element Theory represent phonological categories, which themselves are based on the phonological information that speakers and listeners use to identify morphemes. This information is transferred between speaker and listener via the speech signal, which is defined in acoustic terms. Thus, elements are primarily abstract units of phonological structure, but they also relate directly to some linguistically relevant properties of the acoustic signal. Importantly, if we describe speech sounds in acoustic terms, it becomes possible to identify physical properties which are shared between Cs and Vs. And if they share acoustic properties, we might also claim that they share phonological properties and have features/elements in common. In turn, this allows us to take Assumption 2 seriously. By adopting Assumption 2, Element Theory is able to capture some of the consonant-vowel interactions that standard feature systems cannot express in any natural way.

This paper argues that Assumption 2 is workable and theoretically advantageous. It proposes that Cs and Vs are represented by the same units, the elements of Element Theory. First we look at feature theories, most of which lie between Assumption 1 and Assumption 2. We then turn to Element Theory, which has been developed around Assumption 2. In Element Theory it is claimed that every element can potentially appear in either a C or V representation. For example, the |U| element is interpreted as labiality in Cs and as rounding in Vs; and by recognising |U| as a unit common to both Cs and Vs, it becomes possible to formalise the phonological link between labiality and rounding (Backley and Nasukawa (2009a)). In the latter part of this paper we focus on another consonant-vowel connection, the link between voicing in obstruents and tone in vowels. Our evidence shows that the element |H| is responsible for high tone in vowels and voicelessness in consonants, and that the element |N| represents low tone in vowels and full voicing in obstruents.

2. Feature theories and the consonant-vowel connection

A survey of feature-based theories reveals a tendency for earlier models — i.e. those more closely associated with SPE (Chomsky and Halle (1968)) — to favour Assumption 1. In SPE itself, Cs and Vs are treated as independent categories represented by separate sets of features: Vs have tongue position features like [±high] and [±back] while Cs have place features such as [±anterior] and [±distributed]. So in SPE most V features are irrelevant to Cs, and vice versa. The lack of shared features between Cs and Vs makes SPE’s description of C-V interaction appear arbitrary. For example, the palatalisation rule for [si]→[ji] refers to a trigger [i] that is [±high,−back] and a target [s] that is [±coronal]; and when the rule is applied, [s] undergoes an apparently random change from [±anterior] to a [−anterior] [j].

Various revisions to the SPE feature system have been proposed to address problems in the original model. For example, Odden (2005) follows Assumption 1 yet allows some feature-sharing between Cs and Vs. Most Cs are described by features like [±anterior] and [±coronal], but he also uses vocalic features such as [±high] and [±back] to distinguish
among back Cs (e.g. velar/uvular/pharyngeal). There is, however, no emphasis on natural
classes: e.g. although Odden uses [–high] to describe labials, coronals and pharyngeals (as
well as non-high vowels), he does not suggest that these form a phonologically natural class.

By contrast, Clements and Hume (1995) give priority to Assumption 2, proposing
that C and V representations have the same place node containing the features [labial],
[coronal] and [dorsal]. Furthermore, their features are motivated by C-V interactions and by
the existence of natural classes containing Cs and Vs (e.g. dorsal Cs and back vowels).
Clements and Hume stop short of full consonant-vowel unity, however, as some of their C
features (e.g. [continuant]) are irrelevant to V structures; this is largely a by-product of their
features being articulation-based. Similarly, their V-place aperture node is exclusively a
vocalic property which has no place in C representations.

Now consider Element Theory. To identify features which apply equally to Cs and
Vs, it seems that those features ought to be based on acoustic properties. This is because,
even though certain Cs and Vs maybe distant in articulatory terms, they may still be closely
related in acoustic terms. This is the approach taken in Element Theory, which abandons
articulation-based features in favour of a set of six elements. Each element stands for some
acoustic property of the speech signal, as described in Nasukawa and Backley (2008). In
principle, any element can appear in any syllabic position, so it is free to contribute to C and
V representations. Consequently, a single element has several phonetic interpretations,
depending on its syllabification. In §3 we describe the elements, and show how using the
same element to represent consonantal and vocalic properties allows us to formally express
certain natural classes that cannot easily be captured by traditional features.

3. Introducing the elements

The version of Element Theory used here employs the six elements |I U A H N ?|,
which divide into a vowel set |I U A| and a consonant set |H N ?|. This is a natural division,
as |I U A| represent most vowel contrasts whereas |H N ?| create manner and laryngeal
contrasts in Cs. These groupings are not absolute, however: though |I U A| naturally belong
in Vs, they are regularly found in Cs too; similarly, the consonant elements |H N ?| can
appear in Vs. In principle, there is no restriction on where an element can/cannot occur.

When a vowel element occurs alone in an onset it is interpreted as a glide: |I| gives
us the palatal glide [j] and |U| the labial glide [w], for instance. But when a vowel element
combines with consonant elements, it contributes consonantal place (POA) properties:

(1) nucleon onset

|I| front vowels palatal, coronal POA
|U| rounded vowels labial, velar POA
|A| non-high vowels uvular, pharyngeal POA

By allowing vowel elements in Cs, we easily capture some familiar natural classes. For
example, the |I|-class unites front vowels, palatals and coronals while the |U|-class brings
together rounded vowels, labials and velars (Backley and Nasukawa (2009a)).1 Similarly,
allowing consonant elements in Vs makes it possible to capture other natural patterns. Some of these patterns have been all but ignored in the phonology literature.

A brief explanation of these consonant elements is in order here. In consonants, |?| provides the stopness or oral occlusion which characterises oral and nasal stops. And when no other elements are present, it is interpreted in an onset as [?]. In some languages it can also appear in nuclei, where it creates a laryngealised vowel to give the effect of creaky voice. This is a contrastive vowel property in languages such as Jalapa Mazatec.

The remaining consonant elements |H| and |N| also show consonant-vowel unity; in other words, like other elements they can appear in either Cs or Vs. We can be sure that the same element is present in C and V representations because there exists phonological evidence to support a non-arbitrary connection between the two. In the remainder of this paper we examine the behaviour of |H| and |N| (which is also labelled |L| in some sources). In onsets, these two elements provide laryngeal properties such as voicing and aspiration in obstruents, while in nuclei they create tonal distinctions on vowels. To illustrate the link between laryngeal and tonal properties, we present data which highlights the interaction between high tone and voicelessness; this supports the Element Theory claim that these are both represented by the same element |H|. Our data also highlights a similar interaction between low tone and obstruent voicing; Element Theory accounts for this by claiming that both properties are represented by the same element |N|.

4. Consonant-vowel unity: the laryngeal-tonal connection

The link between laryngeal contrasts and tonal contrasts is illustrated by tonogenesis — the emergence of a tonal distinction in Vs. Typically, Vs develop high tone immediately after voiceless Cs and low tone after voiced Cs. After the tonal distinction establishes itself, the voicing contrast may remain or may disappear. In languages where it survives, a lexical distinction is carried on the C (by VOT) and also on the V (by F0). In those systems where it disappears, however, the lexical distinction is expressed solely by F0 on the vowels. In this latter case, the obstruent voicing distinction neutralises to the unmarked member.

Tonogenesis has been observed in several South Asian languages including Kammu. Originally, Kammu had no lexical tone but did have a voicing contrast in obstruents and sonorant Cs. It also had an aspirated stop series for pronouncing loanwords. The present-day Eastern dialect of Kammu (Svantesson and House (2006)) has a similar system:

<table>
<thead>
<tr>
<th>Obstruents</th>
<th>Sonorants</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless</td>
<td>voiced</td>
</tr>
<tr>
<td>[p t c k]</td>
<td>[m n ŋ l r w j]</td>
</tr>
<tr>
<td>voiced</td>
<td>voiceless</td>
</tr>
<tr>
<td>[b ð j g]</td>
<td>[m n ŋ l r w j]</td>
</tr>
<tr>
<td>aspirated</td>
<td></td>
</tr>
<tr>
<td>[pʰ tʰ cʰ kʰ]</td>
<td></td>
</tr>
</tbody>
</table>
Now compare this system with that of the northern dialect. Northern Kammu also retains an aspirated series for pronouncing loanwords, but the voiced and voiceless obstruents have merged into a single voiceless series. The lexical contrast itself is not lost, however, since it is expressed as a tonal contrast on the following vowel instead:

(4)  

<table>
<thead>
<tr>
<th>Eastern</th>
<th>Northern</th>
</tr>
</thead>
<tbody>
<tr>
<td>taaŋ</td>
<td>táaŋ</td>
</tr>
<tr>
<td>daaŋ</td>
<td>tâaŋ</td>
</tr>
<tr>
<td>tʰaaŋ</td>
<td>tʰâaŋ</td>
</tr>
<tr>
<td>raŋ</td>
<td>râaŋ</td>
</tr>
<tr>
<td>ḍaŋ</td>
<td>râaŋ</td>
</tr>
</tbody>
</table>

Cs followed by high tone in Northern Kammu correspond to voiceless Cs in Eastern Kammu, while Cs followed by low tone in Northern Kammu correspond to voiced Cs in Eastern Kammu. Thus we see a direct association between voicelessness and high tone (the same |H| element) and between voicing and low tone (the same |N| element). Apparently, speakers of Eastern and Northern Kammu have no difficulty in understanding each other. And perception experiments (Svantesson and House (2006)) show that speakers of the two systems are unaware of their phonological differences. Kammu speakers themselves claim that their dialects are distinguished by lexical rather than phonological differences.

Consonant-vowel unity is also apparent in the Songjiang dialect of Chinese, which has the tonal inventory in (5) (the numbers next to each form indicate the pitch contour of the vowel). A total of five tone heights are distinguished, with 1 denoting the lowest tone and 5 the highest. Data are from Halle and Stevens (1991), originally in Bao (1990:64).

(5)  

<table>
<thead>
<tr>
<th>ti</th>
<th>44</th>
<th>‘bottom’</th>
</tr>
</thead>
<tbody>
<tr>
<td>di</td>
<td>22</td>
<td>‘brother’</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>‘emperor’</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>‘field’</td>
</tr>
</tbody>
</table>

The pattern established by (5) is such that words beginning with the voiceless obstruent [t] have tones in the upper register (between 3 and 5), whereas words beginning with a voiced [d] have tones in the lower register (between 1 and 3). This distribution provides further evidence to support the claim that laryngeal and tonal contrasts are phonologically linked. In spite of their physical differences, vocal fold stiffness (i.e. voicelessness) and high tone may be treated as a single feature; in Element Theory terms this is dubbed the |H| element.

A final example of C-V unity comes from Tibetan, where the historical distinction between aspirated and unaspirated stops has become neutralised and replaced by a tone contrast in the present-day Lhasa variety (Silva (2006:299-300), Duanmu (1992)). For instance, the historical form \textit{go} ‘hear’ (voiced stop) has evolved into \textit{khô} (rising low tone). We argue that the voicing of [g] and the low tone of [khô] are different phonetic realisations of the same phonological property. In Element Theory this property is called the |N| element.

How do we account for the facts of Kammu, Chinese and Tibetan? In each case, one set of contrasts can be interpreted in two different ways: as a C property or as a V property. We assume that this is only possible if the same element can appear in onsets and nuclei,
and furthermore, if the acoustic cues to laryngeal and tonal contrasts have properties in common. In §5 we consider evidence for a phonetic link between the two.

5. Phonetic evidence for shared consonant-vowel properties

The feature-based model in Halle and Stevens (1971) unifies laryngeal and tonal properties by analysing pitch change and phonation type in terms of vocal fold tension. The authors claim that when the glottis is narrowly open, the stiffness at the edges of the glottis controls the state/rate of vocal fold vibration. In consonants, especially obstruents, the drop in pressure across the glottis is small, but in vowels and other sonorants the drop in pressure across the glottis is large. In obstruents, stiffness of vocal folds controls the presence or absence of vibrations: slack vocal folds produce voicing and stiff vocal folds produce voicelessness. In sonorants, however, the state of the vocal folds controls rate of vibration (i.e. pitch): slack vocal folds produce low pitch while stiff vocal folds produce high pitch.

Halle and Stevens employ the four laryngeal features \([±\text{spread glottis}], [±\text{constricted glottis}], [±\text{stiff vocal cords}]\) and \([±\text{slack vocal cords}]\), the last two participating actively in Vs and Cs: in nuclei these encode high/low tone and in C positions voicing/voicelessness.

\[
\begin{array}{c|c|c}
\text{Laryngeal features} & [+\text{stiff vocal cords}] & [-\text{stiff vocal cords}] \\
\text{Phonetic realization} & [-\text{slack vocal cords}] & [+\text{slack vocal cords}] \\
\text{in Cs} & \text{voiceless} & \text{voiced} \\
\text{in Vs} & \text{high pitch (F0 up)} & \text{low pitch (F0 down)}
\end{array}
\]

This is consistent with the pattern observed in Kammu, that voiceless Cs correlate with high-pitched sonorants (especially vowels) and voiced Cs with low-pitched sonorants (Yip (1980), Bao (1990), Duanmu (1990), Halle and Stevens (1991)). In fact, Kammu is not an isolated case. Tonogenesis is reported many languages such as Chinese (Yip (1980)), Karen (Burling (1969), Haudricourt (1961), Henderson (1973)), Tibeto-Burman (Matisoff (1972), Mazaudon (1977)), Miao-Yao (Chang (1973)) and Vietnamese (Haudricourt (1954)).

How can Element Theory formalize the patterns described above? Like the Halle and Stevens model, it represents laryngeal categories using units which occur in Cs and Vs. In (7), \(|H|\) stands for aperiodic ‘noise’ energy and \(|N|\) for a low frequency ‘murmur’.

\[
\begin{array}{c|c|c}
\text{Laryngeal categories} & |H| & |N| \\
\text{Phonetic realization} & |H| & |N| \\
\text{in Cs} & \text{noise} & \text{murmur} \\
\text{in Vs} & \text{high pitch (F0 up)} & \text{low pitch (F0 down)}
\end{array}
\]

In consonants \(|H|\) and \(|N|\) represent the phonetic cues ‘noise’ and ‘murmur’ respectively, and in vowels they stand for high and low pitch. Up to this point, there seems little to separate the laryngeal model of Halle and Stevens from that of Element Theory. Yet elements do offer an advantage over orthodox distinctive features. In addition to unifying voice and tonal contrasts, Element Theory captures the link between voicelessness and aspiration, and also
the link between voicing and nasality. It achieves this by assuming that elements combine asymmetrically to create head-dependency relations within the same expression.

When $|H|$ dominates the other elements in an expression, it is known as a headed $|H|$ (written $|H|$). Headedness causes $|H|$ to be interpreted as aspiration. By contrast, a dependent (non-headed) $|H|$ contributes voicelessness. This reflects the idea that the voicelessness inherent in $|H|$ is perceptually stronger in aspirates than non-aspirates, and is consistent with the general asymmetry between heads and dependents in linguistic structure: the properties of a head are expected to be stronger and/or more prominent than those of a dependent.

(8)

| Laryngeal categories | $|H|$ | $|N|$ |
|----------------------|------|------|
| in Cs                | voicelessness | nasality |
|                      | aspiration    | voicing |
| in Vs                | high pitch    | low pitch |
|                      | devoicing     | nasality |

As (8) shows, this dual phonetic interpretation of an element also applies to $|N|$; if $|N|$ is non-headed, it is interpreted as nasality; but when it is headed, it produces obstruent voicing.

The representations in (8) capture a number of useful phonological patterns. Besides formalising the relation between aspiration, voicelessness and high pitch, $|H|$ allows us to express lenition effects in weak positions: e.g. deaspiration $[p^ʰ] \rightarrow [p]$ in English, Swedish, Thai and Bengali (Backley and Nasukawa (2009b)). As for $|N|$, besides formalising the link between obstruent voicing and low pitch, it highlights the nasality-voicing link illustrated by processes such as postnasal voicing (Zoque, Quichua, Kpelle, Yamato Japanese) and velar nasalisation (southern Tohoku Japanese). It also helps explain why voiced Cs are susceptible to nasalisation in nasal harmony languages, and why prenasality tends to lower the pitch of a following vowel (Miao dialects, Chichewa — see Hombert (1978:91)).

6. Conclusion

Using features, many of the above phenomena appear problematic or arbitrary. This is because feature models do not fully exploit the idea that the same features may apply to both C and V representations. By contrast, Element Theory can capture these effects despite using a relatively small set of melodic units. It succeeds because elements are not tied to the articulatory properties of speech, and because they form dependency relations in segments, thus encoding different but related acoustic realisations of the same phonological category.

**************************************************************************

Notes
* This paper was first presented at the Phonology Forum meeting (Kobe) on 29 August 2009. We thank the participants for their helpful comments. This research was partially funded by the Japanese government’s Ministry of Education, Culture, Sports, Science and Technology under grant 19520429.
Elements in the same expression may form head-dependency relations. For example, headed |I| manifests itself as palatality, while non-headed |I| is interpreted as coronality. For further discussion of melodic headedness, see Nasukawa and Backley (2008).

References