NATURE OF PHONOLOGICAL REPRESENTATION

YUKO YOSHIDA
Doshisha University*


Keywords: constraint-based, unification, lexical representation, phonological government, phonological elements

1. Introduction

The recent literature presents a range of theoretical frameworks that reject both phonological derivation and representation; clearly, these deserve careful consideration, assuming that a legitimate representational approach can be called upon for the purposes of comparison. Included in this set of non-derivational theories are Optimality Theory (Prince and Smolensky (1993), McCarthy and Prince (1993), et al.), and Declarative Phonology (Scobbie, Coleman and Bird (1996), Coleman (1998)), both suggesting a form of representation more abstract than that typically exploited within mainstream autosegmental theories. This paper critically evaluates the theory of Declarative Phonology—an

* My gratitude goes to Phillip Backley, and two anonymous reviewers of the English Linguistics for their invaluable and helpful comments, and to Hideki Zamma for practical help. Any defects of this paper are of course mine alone. This paper—and in particular, the material discussed in section 8—is based on the presentation ‘Great Vowel Shift Continues,’ which was given at the Workshop “The Great Vowel Shift and Phonological Theory” at the 16th annual meeting of the English Linguistic Society of Japan.
approach which blurs the distinction between an abstract form of phonological representations and constraints, in order to account for various phonological ‘events’ without recourse to phonological rules and their ordering. After pointing out some of the problems inherent in the Declarative Phonology view, I shall highlight the advantages of recognizing phonetically interpretable phonological representations. The constraint-based view of representational well-formedness supported here contrasts sharply with earlier transformational approaches, in which grammaticality is defined with respect to phonological rules. The possibility of allowing fully interpretable phonological representations relies on a clear understanding of where and how to draw the line between phonology and phonetics, and specifically, on the decision to locate phonology squarely within the grammatical camp—unlike the Declarative approach, which resorts to various dynamic phonetic accounts of phonological events. To achieve this goal, the respective grammatical roles of phonetics must initially be established.

The final section of this paper considers the advantages of employing constraint-driven phonological elements (Charette and Göksel (1994), Kaye (1997)). Eschewing traditional phonological rules and features, these units enjoy realizational autonomy (Kaye, Lowenstamm and Vergnaud (henceforth KLV) (1985, 1990), Harris and Linsey (1995)) and serve to explain fully the cognitive nature of phonology.

2. Core of Declarative Phonology

The proposals put forward by Scobbie, Coleman and Bird (1996) and Coleman (1998) offer an economical and dramatically sharp way of handling lexical representations and phonological constraints, albeit within a largely descriptive focus. In Declarative Phonology, the distinction proposed by Johnson (1988) and Kaplan and Bresnan (1982) is used to separate linguistic objects and the formal theoretical description of those objects. Phonological representation and phonological constraints: every element of the phonology is a partial description. Coleman (1998) pursues the use of logical connectives as the basis for a descriptive language to unify all phonological properties expressed in the form of predicate-argument structures. He transfers the syntactic formalism of Unification Grammar proposed by Shieber (1986) to the domain of phonology, where this approach extends formal language theory to include complex categories (e.g. sets of possibly nested fea-
ture-value pairs) and incorporates insights from type-theories employed in logic and computer science. The relationship between the description and objects is captured using logical connectives. To illustrate, consider the counterfeeding order required within SPE for Welsh mutation, described in Coleman (1998). The extrinsic rule ordering given in (1c) is necessary in order to prevent [t] from alternating with /z/.

(1) a. \([-\text{voice}, -\text{cont}] \rightarrow [+\text{voice}]\) i.e. t → d/X\_Y
b. \([+\text{voice}, -\text{cont}] \rightarrow [+\text{cont}]\) i.e. d → z/X\_Y
c. b<<a (<=strict precedence relation)

In declarative phonology terms (1a) and (1b) are disjoint, since lexical representations cannot interact with one another:

(2) \((t \rightarrow d) \lor (d \rightarrow z)\) (\(\lor=\)logical exclusive ‘or’)

Underspecification is required, so as to render phonetic interpretation possible for those alternants /t-d/ (partially specified for voicing) and /d-z/ (partially specified for continuancy).

An analysis that unifies lexical representations and various constraints through the use of logical connectives provides an elegant approach towards monostratal phonological analysis, thereby abolishing rule ordering. This treatment allows both representations and constraints to be identified as partial (and more or less equal) descriptions, thereby contrasting with the standard transformational approach that postulates lexical representations of linearly ordered segments and allows structural changes to take place via rule application, thus resulting in ‘destruction’ (Coleman’s term for assimilation) and various deletion phenomena. An abstract form of lexical representation merely supplies a series of necessary blocks of information, rendering the linear order of segments redundant. To facilitate this, the notion of dags is employed. These denote the relation ‘is a constituent of,’ and are formed from one or more unordered trees and re-entrancy arcs—that is, nodes dominated by more than one mother (e.g. an ambisyllabic node). From the obligatory head of a dag, relevant features percolate down to optional (dependent) constituents. The following example is taken from Coleman (1998: 206) and represents the English word *cousin*:
Acoustic, rather than articulatory features (Jacobson, Fant and Halle (1952)) are used here to facilitate a cross-classificational distinction (between consonants and vowels) for [grave] (backness/roundness) and [compact] (palatality/velarity). [+grave] describes the relationship between backness and labiality in both vowels and consonants, and [compact] shows the difference between [+compact] for open resonance and palatal and velar consonants and [-compact] for non-open, mid or close resonance and labial and alveolar consonants. Propagation is regarded as an informal interpretation of the effects of feature agreement constraints holding between a feature and its constituent(s). Feature spreading benefits from this approach to encode wide-domain phonological phenomena. A generalization controlling phonotactics within the English rhyme—namely, that only coronal consonants (and never labial or velar consonants) may follow the diphthong /aw/—can be

In a later chapter, Coleman explains that C and V units are redundant since from the feature value, the nodes for consonants and vowels are predictable.
achieved effectively by the approach: as Coleman states, a constraint on rhyme constituency prohibits the nucleus and coda from agreeing for the value of the feature [grave] if the nucleus is an open rising diphthong. Note here that acoustic features serve better to explain and describe the phenomenon than articulatory ones do (cf. gestural model in Browman and Goldstein (1989)), which may only describe the fact in question as 'after sequences of [wide] (for a) and [narrow] (for w) TB (Tongue Back) gestures, [velar] TB (k, g) or [labial] LIPS (p, b, m) gestures may not follow [velar] TB or [labial] LIPS gestures, and [palatal] TT (Tongue Tip) gestures (tʃ, dʒ) may not follow [palatal] TB gestures' (Coleman (1998: 174); the parenthesized material is mine) but does not tell us why this should be the case. The category-valued features [consonantal], [vocalic], and [source], along with the acoustic features [grave] and [compact], are employed in Declarative Phonology to describe cross-categorial relations such as /u/ and /v/ as in solve/solution.

3. The Non-Derivational Approach and Explanatory Power

In terms of offering a generalized treatment of this rhymal phonotactic pattern in English, Coleman’s gravity approach demonstrates some advantages. However, it is still difficult to pin down any real explanation for why a nucleus and a coda within the same rhyme should not agree in value for [grave]. Considering dialectal varieties of English, Harris (1994) points out that systems such as Scottish English and dialects of Northern England do not conform to the Standard English pattern, retaining sequences (aw+labials, velars) found in an earlier stage of the language, e.g. cowp ‘overturn,’ howf ‘burial ground’ and howk ‘to dig into.’ This may well be considered a historical accident.

Coleman rejects theories which may result in overgeneration: the now-abandoned Charm Theory of Government Phonology (GP) (KLV (1985)) should be included as such a theory, since there was no explicit way of restraining the generation of, say, a vowel unattested in a given language. Leaving until later a detailed discussion of the revised theory of phonological elements that effectively ousted Charm Theory, I shall focus here on Coleman’s claim for a highly constraining theory. Aiming to declare, and perhaps to predict, what is really available to generalize, Coleman hinders the nature of phonology, as in his analysis of the Great Vowel Shift. Gravity Dissimilation (GD), discussed for
English rhymal phonotactics, accounts for the Great Vowel Shift (GVS) too, which lends apparent support to the employment of such acoustic features as [grave] and [compact] along with subcategories such as consonantal and vocalic. This time, GD manages to treat the two consecutive vocalic nodes within a nucleus, unlike the previous case where vocalic and consonantal nodes within the rhyme are subject to the constraints.

(4) 1. Open to mid: \( /a:/ \rightarrow \text{diphthongal} /ey/ \) (e.g. name)
   \( /\alpha:/ \rightarrow \text{diphthongal} /aw/ \) (e.g. whole)²

2. Mid to close: \( /e:/ \rightarrow \text{diphthongal} /iy/ \) (e.g. feet)
   \( /o:/ \rightarrow \text{diphthongal} /uw/ \) (e.g. tooth)

3. Close to open: \( /i:/ \rightarrow \text{diphthongal} /iy/ \) (e.g. bite)
   \( /u:/ \rightarrow \text{diphthongal} /aw/ \) (e.g. house)

For the mid to close case, the quality of the off-glide is predictable from the original vowel: when the vowel is front non-open, then the off-glide is front, whereas the back off-glide appears when the vowel is back non-open.

(5) \[
\text{Nucleus} \\
V1 \quad \quad V2 \\
[\text{grave: } \alpha] \\
[\text{height: } \neg \text{open}] \\
[\text{grave: } \alpha] \\
[\text{height: close}] \\
(\neg \text{denotes logical ‘not’})
\]

For the open diphthongs /ay/ and /aw/, the backness of the off-glide follows that of the original vowel. The following two constraints capture the relevant generalization:

(6) a. \[
\text{Nucleus} \\
[\text{grave: } \alpha] \\
V1 \quad \quad V2
\]
   \[
[\text{grave: } \alpha] \\
[\text{height: close}]
\]

b. \[
\text{Nucleus} \\
V1 \quad \quad V2 \\
[\text{grave: } \alpha] \\
[\text{height: open}] \\
[\text{grave: } \alpha] \\
[\text{height: close}]
\]

These constraints together may elegantly describe the generalization of the event, yet they fail to identify any cause of, or motivation for, the

² This data seems to vary from other sources of the GVS: for example in Wells (1982), /\alpha:/ (ME)→ou (Present day).
GVS. This unavoidable weakness originates from the lack of a structured lexical representation, which should identify the very site where the event (phonological operation) occurs, and in most cases the reasons for its occurrence: e.g. an adjacency relationship between the two sounds which are subject to dissimilation or assimilation. Without offering any motivation for why a phonological operation should take place where it does, any explanation must be viewed as implausible.

4. On Lexical Representation

A representational model of phonology has to be defended from Coleman's criticism, regarding to which phonological processes within such a theory involve 'destruction' of representations. In this section, I point out one of the shortcomings of the highly abstract form of lexical representations proposed in Declarative Phonology. The constraint-based view of the representational approach offers a less damaging representation and mode of constraint application. To propose lexical representations to which a set of constraints (universal or language specific) can apply does not necessarily amount to a destructive move. Rather, to leave a terminal node unspecified means that when the word in question appears in isolation, the node has to access a feature or relevant phonetic content from nowhere. The assumption that a variety of phonetic properties for a lexical default specification appear in such cases, which is one possibility that Coleman suggests, should be denied if the purpose of leaving the node unspecified should be to avoid the deletion of any part of the representation. Besides, there is no way of distinguishing effectively between a position lexically fully specified for its phonetic content, and a position unspecified but supplied with a lexical default value. For velar softening in English, Coleman employs a 'lexicalist' approach in which he assumes the lexical representation of electric to be /elektri [k, s, j]/, where the logical connective disjoins [k], [s] and [j]. Perhaps a truly lexicalist hypothesis can only be complete when we assume all the variant forms for a limited set of words such as opaque/opacity are lexicalized and located in separate lexical entries (Kaye (1997)). The processing cost will clearly be much smaller to store each one of electri[k]/electri[s]ity/electri[f]ian as an independent variable (perhaps with semantic association), rather than to posit a lexical structure as such and to unify with many constraints, e.g. to pronounce elektri[k] when this morpheme precedes a zero morpheme, and
to pronounce \textit{elektri}[s] when followed by \textit{-ity}, etc.

5. Roles of Phonology and Phonetics

The position assumed in Declarative Phonology in Coleman (1998) with respect to the phonology/phonetics distinction recasts many phonological processes as phonetic events. Further, Coleman regards as a strength of the Declarative approach the claim that each single 'phonological' phenomenon is open to either phonetic or phonological treatment. With this approach, the distinction between true phonological contrasts and cases of allophonic variation is easily made: the allophonic variation observed between plain and aspirated stops in English arises from the phonetic interpretation of the same phonological object in different structural contexts. This situation should become clearer if contrasted with a language such as Thai, in which aspiration behaves as a phonologically contrastive property, as in \textit{pʰaːa} 'to split' versus \textit{paa} 'forest.' In this instance, the Declarative Phonology approach would likely adopt the convenient position of representing aspiration as a property of phonological, rather than phonetic, representation, thereby contradicting the analysis of English aspiration.

The box (envelope) form of phonetic representation, as is effectively applied to the YorkTalk speech recognition system, describes the temporal relation between the onset and nucleus (and a coda). However, we have to bear in mind that this representation is merely descriptive, and that, despite the dynamic nature of phonetics that Coleman claims, this representation does not provide a suitable foundation on which to base predictions. So, the option of resorting to a phonetic account of something that appears to defy a well-defined phonological treatment, may be construed as no more than a description which avoids the theoretical issue in question.

\begin{align*}
\text{(7) Declarative characterization of /t/ epenthesis in \textit{mince}}
\end{align*}

\begin{table}
\begin{tabular}{c|c c | c c}
\hline
Closure & Friction & = & Closure & Friction \\
Nasality & Non-nasality & & Nasality & Non-nasality \\
\hline
n & s & & n & t & s \\
\hline
\end{tabular}
\end{table}

\textbf{NB:} = denotes non-distinctness

(7) illustrates the non-distinctness of the forms in, for example, \textit{mins}
‘mince’ and min[t]s for speakers who habitually have epenthetic [t]. Coleman successfully describes here the phonetic event in question to demonstrate that a possible phonological analysis is not plausible in which the phonological representation of min[t]s includes a three consonant cluster in the coda. Coleman’s claim is that “phonological representations are not ‘turned into’ phonetic representations by rewrite rules” (p. 176), and that the relation between them is essentially arbitrary although nevertheless systematic. This claimed arbitrariness in his system shows that the phonetics part in Coleman’s terms fails to explain the systematic nature of the relation between phonology and phonetics, and functions as no more than phonetic interpretation, which he criticizes, in derivational theories, e.g. in GP for their nonrestrictive nature of phonetic interpretation from the phonological representation. GP controls the generation of phonological expressions via the postulation of both universal principles and language-specific constraints, thus yielding the most appropriate form as the input to phonetic interpretation. To some extent, the phonetic interpretation in GP must inevitably be language-particular; for example, the phonetic quality of a vowel alternating with zero may vary. In French, a null vowel (an empty nuclear position) alternates with ə, whereas in Tonkawa almost any lexically present vowel may alternate with zero in an appropriate environment illustrated in the next section (see (12)).

What are referred to as assimilatory phenomena in classical terms are analyzed in part as phonetic events in Declarative Phonology, and the purpose of this treatment is to establish phonology as monostratal and to avoid use of rewrite rules for deriving phonetic representations from phonological ones.

(8) tha[t] case → tha[k] case
This type of assimilation is a productive feature in English sound structure, involving a triggering process, a concatenation of two phonological domains. Each lexical item marks out its own phonological domain, indicating the boundaries within which appropriate phonological events are expected to be carried out. Further, Coleman has to give a good account of why thi[l]s] case never appear as thi[l]k] case: the trigger for this assimilation process must be absent from this sequence.

6. The Importance of Skeletal Positions and Lexical Representation

Coleman argues against ordered lexical constraints, relying instead on
simple tree structures to provide syllabic information: a branching rhyme dominated by the optionally branching syllable node. For Coleman (and similarly for other Declarative Phonologists), the lexical representation only serves to provide, so to speak, a mass of blocks to be put together into a well-formed linear structure through unification with other constraints. Clearly, the first question to be addressed is how to distinguish between the respective lexical inputs of *pat* and *tap*. Coleman claims that the linear ordering of segments is entirely predictable in an Onset-Rhyme sequence, and also, that the order of rhymal consonants and vowels is predictable from the following hierarchical structure.

(9) Syllable
    /\           \
   /   \          /   \   
Onset  Rhyme     Nucleus  Coda
   |     |     |     |
   p     a     t

With the aid of linear precedence constraints such as the general ordering principles (nucleus < coda) and (onset < rhyme), the sequence *pat* can be distinguished from *tap*. To distinguish the respective melodic contents of the onset and the coda, at least one further constraint seems necessary—one which unifies with the lexical representation constraint, in order to allocate the relevant features to the rhyme node.

So far so good. However, problematic cases arise for lexical representations without any explicit structure in the discussion of glides and high vowels, and Coleman highlights this through his explanation of the syllabification for a glide and a high vowel in Japanese. In many languages the glides *y/w* are hard to distinguish from their associated high vowels *i/u*. Only in terms of syllabification can the distinction between them be drawn—that is, whether the cross-categorical features percolate to the onset or the nucleus. An alternation between a high vowel and a glide *iu ~ ju*: exploits the simplicity of this analysis, since no structural changes are required. Coleman claims that in VVV sequences in Japanese, one or more Vs may be treated as a glide, as in *yo + u* ‘gets drunk/seasick’ and *huyu* ‘winter.’ This assumption demands to be contradicted, however. Firstly, these glides are not vowels initially. Secondly, it makes little difference whether the string consists of a sequence of vowels or whether it includes lexically spe-
cified consonants. The proposed feature percolation mechanism is misleading as an indicator of lexical contrasts in Japanese, highlighting unpredictability in the lexical structure.

(10) a. you ‘gets drunk/seasick’ iou ‘sulfur’
    b. huyu ‘winter’ huiuti ‘surprise attack’

*huyuti ‘winter land (perhaps for agricultural cultivation)* is not a commonly used word, but is nevertheless possible as a combination of *huyu ‘winter’ ti ‘land’*; crucially, the word’s pronunciation is clearly distinct from that of *huiuti*. If we were to posit another constraint to specify that the features involved should percolate with their category value (consonantal or vocalic), then we might as well set the features in question to the appropriate positions, with temporal ordering and with a specification for which category (consonantal or vocalic) they should attach to.

Coleman’s motivation for treating both front and back high vowels as glides is to reduce the maximum number of vowels in sequences such as *yaei* or *iwau* to three, and then further, to two. The latter suggests that the maximal form of a Japanese syllable should be CV.V.C. In other words, such a treatment serves to avoid vowel sequences that have the appearance of triphthongs morpheme-internally. As Hinds (1986) correctly points out for Japanese, there exists no upper limit on the number of contiguous vowels that may legitimately appear, provided that the generalizations controlling morpheme sequences are respected. Suppose that a new type of fish is found to be easily kept at home for appreciation. Then it would be possible to name it *ieaaoaiuo* ‘domestic blue fish with holy-hock patterns (*ie ‘house,’ ao ‘blue,’ aoi ‘holy-hock pattern,’ uo ‘fish’*) without breaking any of the phonotactic rules of the language. Most of the unattested three-vowel sequences cited by Coleman should be readily available, providing every combination of verb/adjective stem plus inflection and every series of Sino-morpheme is considered (as is done by Coleman (1998) in other attested vowel sequences). This fact lends support to the assumption that all vowel sequences consist lexically of separate nuclei; hence, each vowel is assigned to its own syllable.

The absence of skeletal positions may result in a problem when we come to analyze languages such as French, in which the phonological shape of an article can be sensitive to the presence or absence of a following onset skeletal position. This is observed in the case of *h-aspiré,*
where sensitivity to the existence/non-existence of the position is demonstrated in the alternating forms of the French definite article. The presence of an empty non-nuclear position is argued for by Charette (1991), who offers an explanation for why *h-aspire* words behave exactly like those with initial consonants. Vowel-initial words are preceded by the vowelless article—in other words, the vowel of *la* or *le* is subject to ‘deletion,’ whereas *h-aspire* words, which are phonetically vowel-initial (the initial *h* is not pronounced), are not.

(11)

<table>
<thead>
<tr>
<th>Consonant</th>
<th>Vowel</th>
<th><em>h</em>-aspire</th>
</tr>
</thead>
<tbody>
<tr>
<td>la soupe ‘the soup’</td>
<td>[e] ours ‘the bear’</td>
<td>la housse ‘the dust-cover’</td>
</tr>
<tr>
<td>la Tamise ‘the Thames’</td>
<td>[a] amie ‘the friend’</td>
<td>la hache ‘the axe’</td>
</tr>
<tr>
<td>le bébé ‘the baby’</td>
<td>[a] épée ‘the sword’</td>
<td>le héros ‘the hero’</td>
</tr>
</tbody>
</table>

Partial representations in (12) show the benefit of independent skeletal positions to account for this phenomenon. The presence of an empty onset position found word-initially in an *h-aspire* word (12c) is intervening the two nuclear positions, unlike the vowel sequences (12b).


<table>
<thead>
<tr>
<th>ONON...</th>
<th>ONON...</th>
<th>ONON...</th>
</tr>
</thead>
<tbody>
<tr>
<td>x x x x...</td>
<td>x x ← x...</td>
<td>x x x x...</td>
</tr>
<tr>
<td>l a t a...</td>
<td>l a a...</td>
<td>l a a...</td>
</tr>
</tbody>
</table>

(as in *la Tamise*) (as in *la amie*) (as in *la hache*)

Charette (1991) explains that in French, due to the Obligatory Contour Principle (OCP), the preceding nuclear phonological expression, that of the article *la*, is subject to non-interpretation. A contrast is found in (12c); an *h-aspire* word does have the skeletal position word-initially without dominating a phonological expression. This position, although

---

3 The Obligatory Principle, which Leben (1973) originally proposed for the analysis of tones, is employed here as an extended version to apply to various autosegmental tiers assumed in GP.
inaudible, breaks up the adjacency of the two nuclear positions, those of the article and the word, and so the nuclear position of the article in (12c) is not subject to the OCP unlike the case in (12b). Thus, the independent existence of the skeletal tier, and further, the linear or temporal order in the (lexical) representations are fully motivated.

Coleman employs [+consonantal] to define an onset position where a consonant docks. Without postulating another constraint for percolating only [+cons] and no other features to the onset, such a phenomenon explained above cannot be accounted for. The manner in which these constraints are unified appears desirable in not involving rule (or constraint) ordering; instead however, we find numerous constraints proposed, one after another, to illustrate the lexical properties of a phonological string, which can simply be stated by a representation incorporating linear ordering.

The concern of Coleman to avoid representations with temporal ordering originates from the destruction of phonological representation along with derivation. Below is an argument to demonstrate that the phonological treatment of a ‘deletion’ or ‘epenthesis’ process does not necessarily involve this kind of devastation.

In fact, in GP, various phonological processes such as epenthesis or syncope occur as the manifestation of interpretation/non-interpretation of melodic material. As briefly mentioned above, in GP, French schwa, which is subject to ‘deletion,’ is not a specified vowel; rather, it is analyzed as an empty nuclear position receiving phonetic content in a relevant environment. In Tonkawa, almost any lexically present vowel is subject to ‘hide and seek’ in the same environment (Yoshida (1990)) as French (Charette (1991)). The environment in question is identified via government from the neighboring nuclear position. If the immediately following nucleus government-licenses the nuclear position in question, then the nuclear position is not phonetically interpreted.

---

4 Government Phonology (GP) (KLV (1990)) defines two types of asymmetric binary government (licensing) relation to be held at the projection (skeletal level): head-initial intra-constituent government (ia-c) and head-final inter-constituent government (id, e). Assumed three constituents, O(nset), N(ucleus) and R(yme), contract governing relations:
More precisely, this type of government is called Proper Government (PG). In other words, the position is subject to ‘masking,’ and importantly, an inaudible position does not necessarily indicate a deleted position nor a segment. ‘Masked’ vowels are, technically speaking, the nuclear positions which are subject to the phonological Empty Category Principle (ECP) (KLV (1990), Charette (1991), Kaye (1995)) in GP terms. PG provides one of the ECP related environment, where a(n) (empty) nuclear position subject to PG receives no phonetic interpretation:

\[(13) \quad \text{Proper Government (Kaye (1995: 295))}:\]

\[\alpha \text{ properly governs } \beta \text{ if:}\]

1. \(\alpha\) and \(\beta\) are adjacent on the relevant projection
2. \(\alpha\) is not itself licensed, and
3. no governing domain separates \(\alpha\) from \(\beta\)

\[O \quad N \quad O \quad N\]
\[x \quad x \quad x \quad x\]
\[\beta \leftarrow \rightarrow \alpha\]

A set of actual Tonkawa words shows how this schema operates.

\[(14)\]

\[O \quad N1 \quad O \quad N2 \quad O \quad N3 \quad O \quad N4 \ldots\]
\[x \quad x \quad x \quad x \quad x \quad x \quad x \quad x \quad -x \ldots\]
\[w \quad e \quad n \quad o \quad t \quad o \quad x \quad o \quad o \ldots\]

( i ) a. b. c. d. e.

Intra-constituent government (a-c) Intra-constituent government (d, e)

Instead of an unattested constituent, syllable, in GP, an arbitrary number of repetitions of the pattern \(O\ R\), form a word (see KLV (1990), Charette (1991), Yoshida (1995) for further details of the binary theorem). In GP, the licensing relation is assumed at projection levels, e.g. at the nuclear projection, where solely all the nuclear heads are projected, as in the case of proper government (see (13)).
a. we-ntox-oʔ 'he hoes them' (PG ①)
b. notx-oʔ 'he hoes it' (PG ②)
c. notox 'hoe'

N.B.: prefixes unmarked 3rd. sg. accusative
we- 3rd. sg. accusative
suffix -oʔ 3rd. sg. nominative

The initial nuclear position (N1) is inaccessible for PG and thus remains audible in all the examples (Yoshida (1990)). An affixation makes it possible for the sequences in (14a, b) to contain two nuclei besides the inaccessible initial nuclear position. PG should apply to such sequences. Where PG ② operates after prefixation (14a) and suffixation (14b), the phonetic content of N2 is inaudible, being 'masked' by the effect of this type of government. In all these lexically related forms, a range of vowels in various positions becomes inaudible: the phonetic content or phonological expression is not deleted, but they are effectively 'hidden' by PG. This way, we see that autosegmental representations do not necessarily involve any destructive method. Rather, it becomes possible to see why we should merely hide, rather than delete, the phonological expression: the vowels are always there in the lexical representation to match various forms, as in (14a–c).

7. Syllable and Mora in Japanese

One chapter of Coleman (1998) is devoted to a detailed study of Japanese, and he proposes the canonical maximal syllable structure of Japanese to be:

(15) syllable=n mora [+head, −voice], mora [+head], mora [+head, −consonantal], mora [−head], n mora [+head, −voice]  
(n>0) (nC.CV.V.C.nC)

where mora [−head] represents a geminate consonant and mora nasal, and mora [+head, −voice] represents a mora containing a so-called devoiced high vowel. For example, ʃ̥kuʃ̥k̥o 'blessing' (where ʃ̥ represents a devoiced high vowel) is syllabified as CCVCC.

For the attribute [head] for syllables discussed for Japanese in Chapter 6 (Coleman (1998)), we discover both good and bad news when pitch accent phenomena are taken into account. Pitch accent acts as a key factor when considering the prosodic structure of Japanese. By positing X [+head] (X can be a head), Coleman distinguishes CV pairs,
which may take the headship of the syllables, from CV pairs which only appear as dependents. The merit of this view is to classify a mora, consisting of either a single vowel optionally preceded by a consonant, as distinct from either a so-called moraic nasal [ū] or the initial half of a geminate consonant: an independent mora (which may constitute a syllable by itself) is marked as [+head], while a dependent mora (which only occurs in the coda or rhymal complement position) is marked as [−head]. This makes the desirable prediction that only (C)V sequences may bear pitch accent in Standard Japanese. Also, the possibility of assuming headship of the syllable may be reflected in headship at the word level; it is that position where pitch accent is located.

A less promising aspect of the analysis concerns those instances where both prosodic units mora and syllable are involved: a decision must be made over which unit to employ in order for specific constraints to be unified. More precisely, in accounting for accent assignment, as I explained above, making use of the unit syllable provides an advantage to show which mora, specifically the [+head] mora within the syllable, may bear the accent. Morae, on the other hand, should not be ignored when considering high-pitch assignment since, whether those morae are high-pitched or not, they nevertheless bring about lexical contrasts:

\[(16)\]
\[
\begin{array}{cc}
\text{tsu ku} & \text{‘horsetail’} \\
\text{tsu ku ji} & \text{‘to exert oneself’}
\end{array}
\]
\[
\begin{array}{cc}
\text{(in both forms, the final} & \text{‘period’}
\end{array}
\]
\[
\begin{array}{cc}
\text{fi} & \text{is subject to high vowel devoicing)}
\end{array}
\]

Further, accent location in Japanese is to a large extent predictable (antepenult accent), as is the case with many stress languages where stress assignment is based on metrical structure constructed around nuclear positions. This fact shows that Japanese pitch accent assignment conforms to a metrical structure (Haraguchi (1991)) built on nuclei (Yoshida (1999)). The accent falls onto the third nuclei from the right edge of the word, the head nuclear position of the penultimate head-final foot (Yoshida (1999)). In other words, only nuclei, not onsets or codas, are counted in for accent assignment, and this provides strong support for the hypothesis that those [−head] morae (in Coleman’s terms), that is, mora nasals and geminates, also involve nuclear positions (Yoshida (1999)) (see also S. Yoshida (1991) for the analysis of the ‘mora nasal’ syllabified into onset-nucleus). In addition, devoiced
high vowels, which is originally labeled as [+head] in Coleman’s terms, should not lose its nuclear position (17a). Below are the examples in which three nuclei from the right edge of the words, so to speak the ‘stress window’ contain those morae in question.

(17)  
  a. wa ru gu tʃi (devoiced high vowel in tʃi) ‘slander’  
  b. ka tʃi ka N (mora nasal N) ‘sense of value’  
  c. da ga kø ki (geminate sandwiching an empty nuclear position, ø) ‘percussion instrument’

Those ‘morae’ bearing the [−head] feature value (17b, c) are never accented, but are sensitive to accent assignment. In Standard Japanese, an antepenultimate mora is accented when there is no lexically marked accent in the given phonological word, if that mora is [+head] as in kafibatʃi ‘bowl for sweets.’ Not only a [+head] mora but also a [−head] mora contributes to construction of the metrical organization of a given word, so the two morae to the right of the accented mora (antepenultimate mora) can be [−head] morae as in kaijaiN ‘company employee.’ Having considered these points, it would appear that the definition of ‘syllable’ in (15) is not actually suitable to an account of pitch accent phenomena.

8. On Features—Phonological Elements as Phonological Primes

The binarity of distinctive features in Declarative Phonology serves to show attribute-value. Utilizing acoustic features carries with it the advantage I explained above. However, a further improvement is gained from employing phonological elements which not only encode acoustic information, but also behave as cognitive objects.

The much-criticized charm theory of GP has now been recast as a theory of universal phonological elements operating under phonological constraints. The advantage of positing phonological elements rather than distinctive features stems from their autonomous interpretability, which enables phonological representations to be phonetically interpretable at all levels. As a consequence, a model that locates phonology within the grammar, and supports phonetic interpretation without the need for any independent phonetic representation, is possible for the first time. A distinctive feature is too small a unit to be individually
interpreted, and thus needs to be combined with one or more other features in order to be realizable. A percolated feature in phonology needs to be collapsed into one phonetic representation, and this model contradicts Coleman’s own claim against accommodating a separate phonetic representation, as in transformational theories. The specification of elements (KLV (1985), Charette and Göksel (1996)) is sometimes given in primarily acoustic terms. However, elements are not interpreted as acoustic (or articulatory) events; rather, they act as cognitive objects which perform the grammatical function of encoding lexical contrasts. Phonological elements are the set of A, I, U, H, L and ? (Kaye (1997)), which are realized respectively as the vowels a, i, u, a high tone, a low tone, and a glottal stop. They may exist either individually or in combination as Phonological Expressions (PEs), e.g. the vowel e is the combination of A and I, (AI). Williams (1998) reports from the automatic speech recognition (ASR) view that these subsegmental primes are clearly and consistently expressed in the signal both in their pristine forms and in combination with other primes, so that they can be readily extracted from the acoustic signal.

In GP, possible combinations of elements within a given language restricted by language-specific constraints capture the segment inventories of the language in question. For example, Standard Japanese has a set of five vowels, determined according to the two statements in (18), which in combination yield only a(A), i(I), u(U), e(AI) and o(AU).

(18) Licensing Constraints (LCs) for Standard Japanese

I and U must be headed

All expressions are headed

This is far more economical than a statement such as: Standard Japanese has only a, i, u, e and o, and does not have æ, y, φ, æ, etc.

In Declarative Phonology, the palatalisation of consonants results from the constraint that the CV nodes under the same mother should share features (the same features percolate to the nodes). This approach also implies that even in /pa/, the a-ness should be affecting /p/; whether or not this might be true in acoustic observation, it fails to be identified in phonological terms. There is no language which facilitates a phonological contrast using the consonant p with versus without a-ness. In the case of the difference between palatal and non-palatal —for example, /s/ and /ʃ/— a case of allophonic variation is generated in Japanese, whereas in English this difference is exploited as a contras-
tive property. In GP, palatalisation is viewed as the manifestation of the licensing potential (headedness) of the element I, whereas the element A does not have the same quality: this asymmetric behavior of elements is summed up in the more or less universal LC, ‘I and U must be headed,’ in (18). The head of the onset-rhyme pair, the nuclear position, which is the head of the rhyme, licenses the preceding onset position (KLV (1990)): this universal principle appears in Japanese as follows.

(19) a. h (H), i (I), u (U), a (A)
    b. O R c. O R d. O R
      |    |    |    |
      N   N   N
      |    |    |
      x   x   x   x
H << I H << U H A

[çi] [ϕu] [ha]

Here, the constraints show that the “A element is not a natural head” (Kaye (1994)); the I and U elements act as licensors, thus the licensing potential manifests itself in assimilation, as shown above.

These constraints are shared by most languages: in many languages, I or U are heads, whereas A is not a natural head (Walker (1995)). This constraint can serve as the driving force behind the GVS in English (Yoshida (1998)), at least to explain the motivation of the chain shift of diphthongs. Here I only intend to show that a phonological theory and representation should go beyond the purely descriptive, and have explanatory power with the appropriate restrictive facility.

(20) 1. Open to mid: /a:/ > /æ/ (e.g. name)
      /o:/ > /ʌ/ (e.g. whole)
  2. Mid to close: /e:/ > /i/ (e.g. feet)
       /o:/ > /u/ (e.g. tooth)
  3. Close to open: /i:/ > /a/ (e.g. bite)
       /u:/ > /aʊ/ (e.g. house)

In the beginning of the GVS, the LC for GVS must have been active to work on PEs with A as their heads, which are as follows (detailed data of the vowel shift in (21) and (24) are taken from Lass (1992) and Tateishi (1998); underlined blank part of the bracket denotes headlessness of the PE):
The constraint ‘A is not a natural head’ became active in the 15th-16th centuries, and as a means of avoiding A heads, /a:/ (A) became headless (A_), since there is no other element in its PE to take on headship. /e:/ and /o:/ have the I and U elements respectively to take on the headship from A. Together with the LCs for Present-day English (Kaye (1997)), the motivation and the process of vowel shift from the 15th century to the Modern English period becomes clearer. I only take the position to show why the Present-day English vowel system has changed from its ancestor form, and this study is not intended to show the detailed grammar of each historical period.

Licensing Constraints for Present-day English (Kaye (1997))

LC1 Branching nuclei are headed; non-branching nuclei are headless.

LC2 U and I cannot combine.

LC3 Nothing can license I (I cannot appear as an operator in a headed expression).

Headedness in PEs represents ‘tense’ as opposed to ‘lax’ (expressed by headlessness. The consequences of the LCs are summarized as follows:

LC1 results in a vowel system to have non-branching nuclei, /i:/ (I_), /u:/ (U_), /ɛ:/ (AI_), /ɔ:/ (AU_) and /æ:/ (A_), all of which are headless and thus are lax vowels. Branching nuclei are headed, /i:/ (I), /u:/ (U), /o:/ (UA), /o:\ (AU), et (AI), /a:/ (A).

LC3 rules out /e:/ from the English vowel system.

The lax long vowels which change into their tense counterparts or diphthongs are summarized below, with illustration of the active LCs at the relevant points.

\[
\begin{array}{cccc}
\text{ME} & \text{16th c.} & \text{17th c.} & \text{Present day} \\
\hline
\text{GVS LC} & \text{LC1} & \text{LC1&GVS LC} & \text{LC3} \\
\hline
/a:/ & *(A) \rightarrow /æ:/ & *(A_) \rightarrow /ɛ:/ & *(IA_\_\_) \rightarrow /εt/ \\
/ɛ:/ & \rightarrow /ɛ:/ & *(IA_\_\_\_) \rightarrow /ɛ:/ & *(IA) \rightarrow /i:/ \\
/ɔ:/ & \rightarrow /ɔ:/ & *(UA_\_\_\_) \rightarrow /o:/ & *(UA) \rightarrow /ou/ \\
\hline
\end{array}
\]

/a/ underwent diphthongization to et, being subject to GVS-LC as I explained above, and LC1. A should not be the head of the PE and to
maintain the length (since vowel length creates another dimension for lexical contrasts) the most likely element I (see LC3: Nothing can license I) takes up the headship. $\varepsilon$: becomes $i$: via the form $e$: owing to the conditions that 1) it has to maintain length and thus has to be headed, and 2) once A takes the headship, 3) then I element should not be licensed, and thus has to take the headship. $\varphi$: follows the LC1 to be a headed expression, and in the choice of the head, (A should not be a head, being subject to GVS LC) U turns out as the head of the PE.

Suffice it to show the reason why the GVS took place in English, and that the account is derivable from the universal constraints proposed for phonological expressions; no further explanation will be given here. At least, however, this discussion offers support for the claim that phonological elements with proper constraints to restrict over-generation have a clear advantage over the largely descriptive properties of distinctive features.

9. Conclusion

Declarative Phonology is an innovative theory, however, great improvement is awaited on the nature of constraints, which should capture more of the universal nature and explanatory power of a phonological grammar than claimed descriptiveness. To sum up, the good point of the declarative approach is that the formalism is elegant, based on unification of constraints by logical connectives, even the lexical representation as one of the constraints. Declarative Phonology nonetheless treats the languages as where various language specific constraints work, only within the descriptive capacity. We should look forward to detailed in-depth studies of individual languages, including a further study of the languages which Coleman (1998) deals with, such as Japanese, which would enrich and widen the future of this theory.

REFERENCES

Charette, Monik and Asli Göksel (1994) “Vowel Harmony and Switching in


Shieber, Stuart M. (1986) An Introduction to Unification-Based Approaches to Grammar, Center for the Study of Language and Information, Stanford, CA.


Institute for Language and Culture
Doshisha University
1-3 Miyakodani, Tatara
Kyo-Tanabe City
Kyoto 610-0394
e-mail: yuyoshid@mail.doshisha.ac.jp