Keley-i Consonant Assimilation

Diana Archangeli

1 Introduction

Keley-i, a Philippine language, has two rules assimilating consonants across a vowel. [1] Such rules might be taken as evidence against the Morphemic Tier Hypothesis (MTH) and against the Locality Condition (LC). The MTH states

(1) Morphemic Tier Hypothesis (MTH)

If and only if two segments are members of separate morphemes are those two segments aligned in separate phonological tiers.

The Keley-i data suggest that the MTH does not hold universally because consonants assimilate across vowels, which has been taken as evidence for two segmental planes in order to prevent the crossing of association lines:

(2)

\[ x \]

\[ C \]

\[ V \]

\[ C \]

\[ y \]

The data also create problems for the Locality Condition:

(3) Locality Condition (LC)

A phonological rule is applicable only if the target and trigger are adjacent.

The consonant features assimilate across an intervening vowel: the target and trigger, being skeletal slots, are not adjacent.

I suggest here that adopting the feature hierarchy as proposed in Archangeli and Pulleyblank (1986) (which is a modification of Clements 1985) combined with underspecification theory (Archangeli 1984, Pulleyblank 1986, Archangeli and Pulleyblank 1986) allows an analysis of the Keley-i data which

---

1. I would like to acknowledge the contributions Doug Pulleyblank made to the theory developed here, although he is not responsible for the content of this article. I also would like to thank Dick Demers, Ken Hill, and Dick Oerhle for discussion of the analysis of Keley-i presented in this paper. My work was in part supported by NEH Humanities Summer Stipend FT-27533-86, for which I am grateful.
permits maintaining the MTH and the LC. A further result is that the Spreading Hypothesis is maintained as well, thus supporting the hypothesis that phonological assimilation is formally expressed in one manner only, namely by insertion of association lines, and not by feature copy rules. (See Hayes 1986, Archangeli and Pulleyblank 1986.)

(4) Spreading Hypothesis

Phonological assimilation is expressed only by rules adding association lines.

The discussion is organized as follows. First, the feature hierarchy and the theory of underspecification are briefly outlined. I then present a partial analysis of the Keley-i data. The analysis consists of a syncope rule and some rules of consonant assimilation. Finally, I return to the problems that Keley-i presents for the MTH and the LC and propose that the relevant Keley-i data are not only in accordance with the MTH and the LC but predicted by the interaction of the two sub-theories, the Feature Hierarchy and Underspecification.

2 The Two Sub-theories

2.1 The Feature Hierarchy

The feature hierarchy is an organization imposed upon the distinctive features by Universal Grammar, arranging the features into a tree-like structure with nodes dominating features and/or other nodes. Nodes are indicated by "o" and features by "F".

(5)

```
Root Node
  O

Laryngeal Node
  F

Supra-Laryngeal Node
  F

Place Node
  F

Secondary Place Node

F F F
```

Secondary Place features include [high, low, back, round, ATR]; Place features include [anterior, coronal, distributed]; Supra-Laryngeal features include [nasal, lateral, continuant, sonorant, strident]; Laryngeal features
include [spread, constricted]; Tone features include [upper, raised]. Note that the node names are in part mnemonic. For example, although tone features are laryngeal, they do not fall under the laryngeal node because of the way these features behave in the phonologies of languages.

The original motivation for this theory comes from the desire to express natural classes among features, that is, groups of features which function in a unified manner. For example, in Keley-i, the nasal /n/ assimilates to the place of articulation of a following consonant, a process expressed in this theory as spreading of the Place Node, rather than as simultaneous assimilation of several unrelated features. See Clements (1985), Archangeli and Pulleyblank (1986) for motivation and for further details.

In Archangeli and Pulleyblank (1986) it is proposed that in a fully specified representation, a Rime Node (which also marks a syllable head) dominates a Macro-Node (which corresponds to the skeletal slot), and the Macro-Node dominates all other nodes and features. Macro Nodes dominated by Rime Nodes correspond to the V-slots of McCarthy (1979), Clements and Keyser (1983), etc., and Macro Nodes which are not dominated by Rime Nodes correspond to the C-slots.

\[
\text{(6)}
\]

This allows representations of skeletal templates without reference to "C" and "V" (see Levin 1985). For example, a CVCC template is represented underlyingly as $o_M o_R o_M 0_M$, where "o_M" represents a Macro Node and "o_R" a Rime Node.

It is important to point out that the Rime Node posited here is part of syllable structure. Its role in syllable structure is equivalent to that of the presyllabified skeletal position in Levin (1985). The Rime Node is not a device to express [+syllabic] without naming the feature.

2.2 Underspecification

Underspecification theory proposes (i) that features which are not distinctive in a language or for a particular segment in a language are not represented until late in the derivation (this can result in some segments having no feature specifications whatsoever); and (ii) that if a feature is included in underlying representation, only one value (+ or -) is represented at that level, with the other value supplied during the course of the derivation. See Archangeli (1984), Pulleyblank (1986), Archangeli and Pulleyblank (1986). The value which is represented is determined by language universals (corresponding to markedness of particular feature combinations) unless there is language-particular evidence to the contrary. Thus the same sound is not necessarily specified in the same manner in all languages. For
example, the following have been proposed as the completely unspecified vowel: /+/ in Korean (Sohn 1986), /e/ in Spanish (Harris 1980) and in Ainu (Ito 1984), /a/ in Klamath (Levin 1985) and in Nyangumarta (Sharp 1986), and /ll/ in Yokuts (Archangeli 1984) and in Yoruba (Archangeli and Pulleyblank 1986).

Combining underspecification with the hierarchical representation raises the question of the status of nodes which do not dominate any features. In Archangeli and Pulleyblank (1986) it is argued that nodes exist by virtue of dominating features — in general, if there are no features, there is no node either. The one exception to this is the maximal node, i.e. a Rime Node (for vowels) or a Macro Node (for consonants). It is necessary to permit featureless Rime and Macro Nodes as these correspond roughly to the partially syllabified skeleton, and there is good evidence that skeletal positions exist without features, including the feature "syllabic" (cf. McCarthy 1979, Clements and Keyser 1983, Archangeli 1983, Marlett and Stemberger 1983, Levin 1985, etc.).

3 Keley-i

3.1 The Underlying Specification of Vowels

From the discussion in Hohulin and Kenstowicz (1979), Keley-i has five vowels, /i, e, a, o, u/. These five vowels behave asymmetrically with respect to a rule of syncope in the "two-side-open" environment, VC.CV: only the vowel /e/ deletes by this rule. This is seen when the forms with the infix -in-, marking focus on the direct object in the past tense, are compared with forms with the prefix ?in-, marking focus on the instrument in the past tense.

2. For example, if a segment is not specified for any Secondary Place features, then that segment has no Secondary Place Node. If a segment has no Secondary Place features and no Place features, that segment has neither a Place Node nor a Secondary Place Node. On the other hand, if a segment has Secondary Place features, but no Place features, it has both Nodes, because the Place Node dominates the Secondary Place Node.

3. Ken Hill (pc) suggested that this might be a rule of epenthesis rather than deletion. I have rejected this hypothesis because of the distribution of this vowel — it sometimes surfaces in the middle of what would otherwise be a true geminate. If this were analyzed as a rule of epenthesis, then it would be necessary to explain why Keley-i, unlike other languages, does not maintain the inalterability of geminates. See Hayes (1986), Steriade and Schein (1986) on the inalterability of geminates.

4. As Hohulin and Kenstowicz (1979) point out, the morphology in Keley-i marks tense and focus by a single process, not by two distinguishable processes.
In (7a–c), the first root vowel is /a/, /i/, or /u/ and syncope does not apply. In (7d), the first root vowel is /e/ and syncope applies. This sort of asymmetry may, of course, be accounted for by positing a syncope rule deleting [-high, -low, -round] vowels in the VC.CV environment. However, such an account implies that the feature specification [-low] is critical in Keley-i, and it is not: the four vowels /i, e, o, u/ require only specifications for [high] and [round] to be distinguished from each other, and /a/ may be distinguished from the four others by [+low]. [-low] is not crucial. (I use the notation /a/ etc. to indicate the underlying features specifications which surface as [a] etc. provided only redundancy rules apply to the segment. Thus /a/ stands for a segment marked only by [+low], and so forth.)

Furthermore, as proposed in Pulleyblank (in press), asymmetrical behavior of one segment suggests that segment as a candidate for the least specified sound of its type. Here, /e/ is behaving asymmetrically. If /e/ is completely unspecified, the syncope rule simply deletes a featureless Rime Node in the correct environment:[5]

5. In the Syncope rule, "o_r" stands for a Rime Node and o_M stands for a Macro Node. The rule is roughly equivalent to

\[ V \rightarrow \emptyset / V C \_\_ C V \]

I expect that syncope is best expressed in terms of syllabification but I have not yet seen a satisfying way of expressing such rules. Thus I use the linear representation.
SYNCOPE: \( O_R \) \( \rightarrow \) \( \emptyset \) / \( O_R \) \( O_M \) \( O_M \) \( O_R \)
The vowels themselves are represented in underlying representation without the specifications [-high], [-back], and [-low] -- and without specifications for any other feature:[6]

3.2 Medial Gemination

The future with subject focus is marked by the prefix ?um-: in these examples the future subject focus contrasts with the past instrumental focus, marked by the prefix ?in-. Here, the prefix is all that marks the difference.

However, if the verb root has no medial cluster, in general the second root consonant geminates in the non-past tenses. Thus, the forms are distinguished both by the prefix and by the presence or absence of gemination.

6. The absent feature specifications are supplied by redundancy rules. See the references cited, particularly Archangeli (1984) and Archangeli and Pulleyblank (1986).
The alternations between (11-12) are straightforward if we assume a Macro Node (C slot) is inserted between the two syllables of the stem. A regressive assimilation rule then accounts for the geminate. (The notation "o_M" indicates a Macro Node and "$" a syllable.)[7]

(13)

MEDIAL GEMINATION: \( \emptyset \rightarrow o_M / [\text{fut} \, \$_{\text{fut}} \, _{\text{fut}} \text{ }] \)

(14)

REGRESSIVE ASSIMILATION: \( o \quad o \quad \text{Macro Node} \)
\[ \quad \quad \quad \quad \quad o \quad \text{Root Node} \]

Let us compare the derivations of a triconsonantal and a quadricsonantal root to see the effects of these two rules:[8]

(15)

STEP 1: AFFIXATION and FUTURE GEMINATION

a. ?um-ba-C-yu?  \quad b. ?um-dun-C-tuk

7. In the parametric rule formalism proposed in Archangeli and Pulleyblank (1986), the assimilation rules, Regressive Assimilation and Progressive Assimilation (below), need not mention the Macro Node. These rules insert structure (an association line) with the Root Node as the argument. Universal Grammar states that the Root Node may dock only in the Macro Node, an organization that is not manipulatable by phonological rules. Consequently this information is not necessary in the expression of the rule.

8. Macro Nodes and Rime Nodes are represented as Cs and Vs respectively in these figures.
**STEP 2: REGRESSIVE ASSIMILATION**

<table>
<thead>
<tr>
<th>Macro N</th>
<th>C V C - C V - C - C V C</th>
<th>Macro N</th>
<th>C V C - C V C - C - C V C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root N</td>
<td>0 0 0 0 0 0 0 0 0</td>
<td>Root N</td>
<td>0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>etc.</td>
<td>? u m b a y u ?</td>
<td>etc.</td>
<td>? u m d u n t u k</td>
</tr>
</tbody>
</table>

From (15), it looks like we should get *[?umdunttuk] alongside the correct [?umbayyu?]. However, the syllable structure of Keley-i allows only CV and CVC syllables. Thus the first t of ?umdunttuk is unsyllabified and so does not surface, which is the correct result.

Consider now the future construction of verb roots which are triconsonantal and have /e/ in the first syllable. The contrasts with the past are repeated.

(16)

<table>
<thead>
<tr>
<th>future, subj focus</th>
<th>past, instr focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ?um-bebhat ?im-bhat</td>
<td>'cut rattan'</td>
</tr>
<tr>
<td>b. ?um-deek ?in-d?ek</td>
<td>'accuse'</td>
</tr>
<tr>
<td>c. ?um-teppen ?in-tpen</td>
<td>'measure'</td>
</tr>
<tr>
<td>d. ?um-kekbet ?in-kbet</td>
<td>'scratch'</td>
</tr>
</tbody>
</table>

If Medial Gemination (13) applies to these verbs, a Macro Node is inserted after the first syllable (phonemes are superimposed on skeletal slots in (17)):

(17)

| a. ?um-be-C-hat | c. ?um-te-C-pen |
| b. ?um-de-C-ek | d. ?um-ke-C-bet |

What must not happen is for Regressive Assimilation to spread the second root consonant to the empty Macro Node. This can be prevented by first spreading rightwards -- across the vowel /e/. Without underspecification, this results in a Crossing Constraint violation. With underspecification, there is no such violation. In (18), the first column gives the derivation with underspecification and the second the derivation with full specification.

(18)

**STEP 1: AFFIXATION AND FUTURE GEMINATION**

<table>
<thead>
<tr>
<th>C V C - C V - C - C V C</th>
<th>C V C - C V - C - C V C</th>
</tr>
</thead>
<tbody>
<tr>
<td>? u m b a t</td>
<td>? u m b e t</td>
</tr>
</tbody>
</table>

- 8 -
STEP 2: PROGRESSIVE ASSIMILATION

\[
\begin{array}{c}
\text{CV} \text{C - CV - C - CV} \\
\text{umb} \quad \text{hat} \\
\end{array}
\quad \quad \quad \quad \quad \quad \quad
\begin{array}{c}
\text{CV} \text{C - CV - C - CV} \\
\text{umb} \quad \text{be} \quad \text{hat} \\
\end{array}
\]

STEP 3: REGRESSIVE ASSIMILATION: n/a

At step 2 in the second column we see the Crossing Constraint violation. Progressive Assimilation spreads all features: it is a case of Root Node spread. The Root Node spreads across the unspecified Rime Node.[9]

(19)

PROGRESSIVE ASSIMILATION:

\[
\begin{array}{c}
\text{Macro Node} \\
\text{Root Node} \\
\end{array}
\]

If Progressive Assimilation applies, Regressive Assimilation is not applicable, as there is no longer a free target.[10]

Any vowel may occur in the first root syllable but only /e/ is invisible for the purposes of consonant spread in the future construction (see the forms in (16) and the derivation in (18a)). The other vowels, all of which have some feature specifications, block the rightward spread of the root consonant (as seen by the forms in (12) and the derivations in (15)). In all of these cases the first root consonant does not spread across the vowel. Spread across the vowel occurs only when the first root vowel is /e/.

9. The application of these assimilation rules seems to be an Elsewhere relationship in that both target a free Macro Node and if one applies, the other does not. However, the Elsewhere Condition does not order these rules appropriately as there is no obvious subset relationship between the two structural descriptions.

10. Dick Oerhle (pc) suggested an alternative hypothesis, with two different assumptions: (i) instead of inserting a single C-slot for Medial Gemination, a CVCCVC template is inserted (a proposal suggested by Ken Hill as well), and (ii) the root melodies associate to the skeleton from right to left rather than from left to right. The immediately obvious advantage is that the two assimilation rules are reduced to a single leftward spread rule. The disadvantage is that association proceeds opposite to the direction expected given the Universal Association Convention (see Pulleyblank 1986). There are a few cases in the literature where it is argued that directionality is a parameter of the UAC, for example Archangeli and Pulleyblank (1986) and Clark (1983). This proposal is one that I am exploring -- whether it is ultimately correct or not does not affect the arguments here for the complete underspecification of the vowel /e/.
3.3 The Contrastive Identification Morphology

There are two contrastive identification prefixes presented in Kenstowicz and Kisseberth (1979), me- and ne-, giving the meaning of "It was John who Verbed..." The first occurs with the future tense and the second with the past. What is particularly interesting about these morphemes is that they induce nasalization on the initial consonant of the verb root itself. Compare the instrumental focus past with the contrastive identification past:

(20)

<table>
<thead>
<tr>
<th>cont id</th>
<th>instru foc</th>
</tr>
</thead>
<tbody>
<tr>
<td>ne-nuntuk</td>
<td>?in-duntuk</td>
</tr>
<tr>
<td>ne-mayu?</td>
<td>?im-bayu?</td>
</tr>
<tr>
<td>ne-gubat</td>
<td>?ig-gubat</td>
</tr>
<tr>
<td>ne-mdug</td>
<td>?im-pedug</td>
</tr>
<tr>
<td>ne-gbet</td>
<td>?ig-kebet</td>
</tr>
</tbody>
</table>

The CV pattern of the forms in (20d) are different from the others because of the application of Syncope. Derivations follow. Syncope applies first (although it could also follow Nasal Assimilation - the order is not crucial), deleting the first root vowel if it is /e/:

(21)

STEP 1: AFFIXATION

a. CV - C V C V C
b. CV - C V C V C

STEP 2: SYNCOPE

a. ne-bayu?
   ne pedug

STEP 3: ASSIMILATION

a. CV - C V C V C
   : :
   +nas
b. CV - C V C V C
   +nas

The surface forms (21a) nemayu? and (21b) nemdug result.

At the third step in the derivation, [+nasal] spreads rightwards onto
the following consonant, ignoring the intervening, featureless vowel. The feature [\(+\text{nasal}\)] combines with the Place specifications to create [m] from /p, b, /g/ from /k, g/, etc. The correct surface forms are derived.\[11\]

In these cases, [\(+\text{nasal}\)] spreads to the initial consonant of the root, regardless of how that consonant is specified. Initial /h/ and /?/, which have no nodes except Root and Laryngeal, are targets for this rule as are segments specified with some oral features, like /p/, /k/, etc., cf. ne-nulat vs. ?in-hulat 'cover'; ne-oinum vs. ?in-?inum 'drink'. In the formalism of Archangeli and Pulleyblank (1986), this is a maximal rule, with the Macro Node as its target, generating nodes where necessary to fill in gaps between [\(+\text{nasal}\)] and the Macro Node.\[12\]

4 The Two "Problems"

4.1 The Morphemic Tier Hypothesis

The MTH (recall (1)) permits two segments on separate planes if and only if they are in separate morphemes. This immediately raises a question because there are a number of cases adduced in the literature in which some vowel spreads across some consonant (for example Ainu (Ito 1984), Yokuts (Archangeli 1985), Javanese (Kenstowicz 1986), Rotuman (McCarthy 1986)). A combination of the feature hierarchy and underspecification explains these effects: in many languages vowels can be distinguished only by Secondary Place features while consonants frequently do not require specifications of these features. Hence, the Secondary Place Node or some Secondary Place feature may spread from vowel to vowel across consonants without violating the Crossing Constraint and without violating the MTH. (In (22), theVs and Cs are included only for expository purposes. They have no theoretical significance.)

\[\text{-----------}\]

11. I am assuming here that [+voice] is specified on voiced obstruents, and that voice specifications for unspecified obstruents and for all sonorants are supplied by late redundancy rules.

12. Some rule or rules convert [+\text{nasal}, +\text{constricted}] to [\eta] and [+\text{nasal}, +\text{spread}] to [\eta]. As shown in Archangeli (1986) this "readjustment" must apply prior to a rule assimilating place features from a consonant to a preceding underspecified nasal. Effects of this rule are seen in figures (7) and (11), among others.
As seen in (22a), individual vowel features can spread from Secondary Place Node to Secondary Place Node without violating the Crossing Constraint and, as seen in (22b), Secondary Place Nodes can also be shared across an intervening Place Node without violating the Crossing Constraint.

Inspection of (22) suggests that consonants cannot spread across vowels without creating a Crossing Constraint violation. However, the combination of underspecification and the feature hierarchy predicts that consonants do spread across vowels in a very limited environment:

(23) A Prediction

If a rule spreads a consonant across a tautomorphemic vowel, the rule applies if and only if the vowel is completely unspecified.

This is precisely the situation we find in Keley-i. The consonants spread across only the vowel /e/, i.e. the vowel already shown to be the completely unspecified vowel. There is therefore no MTH violation.

4.2 The Locality Condition

The explanation above of why there is no MTH violation indicates that there is a violation of the LC: in a "CVC" string: the two Cs are not adjacent and so spread should not obtain between them if the LC is a universal. I show here that in the particular configuration in which the intervening vowel is unspecified, the two flanking Cs are adjacent. Consequently there is no LC violation. The MTH is not maintained at the cost of the LC. To understand the argument, let us first reconsider the two ways in which vowels are adjacent, illustrated schematically in (22a,b).

In (22a) the feature F spreads from one vowel to the next, as one might see in a harmony process. Focussing more narrowly, we could also say that the feature F spreads from one Secondary Place Node to the next. (The Place Node is irrelevant to the rule. It is included for expositional purposes.)
The Secondary Place nodes are clearly adjacent, since the intervening consonant has no Secondary Place features and hence no Secondary Place Node. The LC is not violated.

In (22b), we see that the entire Secondary Place Node has spread from one vowel to the next, a case of total assimilation. In such a case, the target node is typically the maximal node, i.e. the Rime Node. (The nodes between the Rime Node and the Secondary Place Node are irrelevant to the rule but are included for expositional purposes. The nodes intervening between the spreading Secondary Place Node and the leftmost Rime Node are inserted by operation of the rule.)

By inspecting the representation, we see that the two Rime Nodes are adjacent since the intervening consonant has no Rime Node, only a Macro Node. Again, there is no LC violation.

The Keley-i cases are of particular interest because, unlike the above examples, in Keley-i we find assimilation of a consonantal feature [+nasal] and of the consonantal Root Node across a vowel. However, the intervening vowel arguably has no features. Thus, its only node is the Rime Node. The two Macro Nodes are adjacent and there is no LC violation.[13]

13. Rudimentary syllable structure is included in these figures to maintain the linear order of the Rime Node (which has no Macro Node) and the two Macro Nodes. See Archangeli and Pulleyblank (1986) for more detail.
In order to maintain the Morphemic Tier Hypothesis and the Locality Condition in the face of the Keley-i data, we have adopted underspecification theory and the feature hierarchy, augmented by Rime and Macro Nodes. The facet of the interaction of these two phonological sub-theories which permits us to maintain the MTH and the LC is the claim that non-maximal nodes (i.e. nodes other than the Rime and Macro Nodes) do not exist unless they dominate at least one feature. In that all of these theoretical proposals are argued for independently of the data considered here, Keley-i lends support to these theories and to the proposal that all assimilation is the result of the addition of autosegmental association lines.
References


Kenstowicz, Michael and Charles Kisseberth. 1979. Generative Phonology, Academic Pr, NYC.


