The sonority hierarchy and NO LONG VOWEL: Theoretical implications

0 Introduction: The problem

Many researchers have observed that long vowels are more marked than short vowels (Paradis 1988, Kaye 1989, Prince and Smolensky 1993:§10.3.4, Marotta and Savoia 1994, Rosenthal 1994, Sherer 1994:ch. 2, Hammond 1997:9), a fact encoded in the constraint NO LONG VOWEL (‘NLV’) in Optimality Theory. The precise formulation of the restriction against bimoraic vowels must be considered carefully, however, and in this paper I show that incorporation of the sonority hierarchy into OT has major implications for the nature and content of NLV. Following Zec (1995), I assume that moraicity is mediated through sonority. Under this approach, the generalization that less sonorous segments in a language may be moraic only if more sonorous ones are, derives from the sonority hierarchy. I present an implementation of this principle in OT and discuss benefits and consequences of such an analysis. I show that NLV introduces an apparent redundancy into the set of constraints, Con, in that two constraints govern the moraic status of vowels, but that this is a necessary move.


- May be encoded in Optimality Theory as the constraint NO LONG VOWEL (‘NLV’)

Goals: To show that the precise formulation of the restriction against bimoraic vowels must be considered carefully, and that incorporation of the sonority hierarchy into OT has major implications for the nature and content of NLV.
Zec (1995): Moraicity is mediated through sonority. Under this approach, the generalization that less sonorous segments in a language may be moraic only if more sonorous ones are, derives from the sonority hierarchy.

Theoretical contribution: I present an implementation of this principle in OT and discuss benefits and consequences of such an analysis. I show that the proper formulation of NLV introduces an apparent redundancy into the set of constraints, \( Con \), but a necessary one.

1 Sonority, moraicity and the sonority hierarchy in OT.

In (1) on the handout, we see a sonority hierarchy of the sort proposed by Clements (1990), where principal divisions are determined by major class features. This ordering encodes the greater or lesser ability of a segment to form a syllable peak (that is, nucleus) or margin (that is, onset or coda):

(1) The sonority hierarchy: (based on Clements 1990)

\[
\begin{array}{cccccc}
\text{More sonorous} & \leftarrow & \text{Less sonorous} \\
\text{vowels} & \text{glides} & \text{liquids} & \text{nasals} & \text{obstruents}
\end{array}
\]

Zec, taking only the features [consonantal] and [sonorant], arrives at the simplified classification in (2), to which additional features may be added to derive finer-grained distinctions in sonority.

(2) Major class features of segments (from Zec 1995)

\[
\begin{array}{cccc}
\text{[cons]} & \text{[son]} \\
vowels & - & + \\
sonorants & + & + \\
obstruents & + & -
\end{array}
\]
The first major break in sonority separates segments that are not consonants from those that are. From with the class of consonants, the second major break in sonority separates those segments that are [+sonorant] from those that are [-sonorant], that is, the obstruents.

From this, and following a moraic theory of the syllable in which the mora is the primitive subsyllabic constituent (Hyman, Hayes, McCarthy and Prince, Zec, others), Zec argues that the best way to understand the moraicity of segments is through their sonority. That is, the fact that more sonorous segments occupy the peak position of the syllable (= nucleus) and less sonorous ones occur toward the syllable margins (= onset, coda), Zec argues, is because sonority constraints are imposed directly on prosodic structure, which in turn immediately affect the segments that the moras dominate. Shown in (3) are the major sonority classes, and, hence, classes of moraic segments that result from Zec’s system:

(3) Major sonority classes (and, hence, classes of moraic segments): (from Zec 1995)

a. [-cons] = vowels (only vowels are moraic; as in Khalkha Mongolian, Yidiñ, Spanish)
b. [+son] = vowels and sonorants (vowels and sonorants are moraic; as in Lithuanian and Tiv, proto-Romance)
c. — = all segments (all segments are moraic; as in English, certain Arabic dialects, Latin)

The sets of moraic segments that result from this approach require at least moraic vowels. Additionally, the generalization that less sonorous segments in a language may be moraic only if more sonorous ones are, derives from the sonority hierarchy. That is, there is no way to select only the set of obstruents as moraic without sonorants and vowels being moraic as well, and there is no way to select sonorant consonants as moraic
without the vowels also being moraic, though obstruents may be (as in Lithuanian, Tiv and proto-Romance.)

Applying moraic theory to Clements’ version of sonority hierarchy, and thus maintaining Zec’s insight into the relationship of sonority to moraicity, the universal sonority hierarchy may be captured in OT as in (4). We invert the hierarchy in (1) and formulate a series of constraints that militate against the moraic status of each class of segment (see Prince and Smolensky 1993, Sherer 1994:ch. 2, Hammond 1997). This encodes the lesser ability of an obstruent in relation to a vowel to bear a mora (and thus potentially be the nucleus of a syllable), a fact that is reflected in the higher ranking of *Obstruent\(\mu\) in relation to *Vowel\(\mu\):

(4) One conception of the sonority hierarchy in OT:

\[
*O(bstruent)_\mu >> *N(asal)_\mu >> *L(iquid)_\mu >> *G(lide)_\mu >> *V(owel)_\mu
\]

(see also Prince and Smolensky 1993, Sherer 1994:ch. 2, Hammond 1997)

The fact that vowels may be syllable peaks in all languages derives from the low ranking of *V\(\mu\), whose violation is always minimal.
(5) Some data that follow nicely from incorporation of sonority hierarchy into OT:

(i) **Historical change**: progressive simplification of consonantal system of Late Spoken Latin / proto-Romance that mirrors the sonority hierarchy (Holt 1997):

(A) **Geminate obstruents simplified first**
(before appearance of first documents written in Romance):

/\textipa{pp}/ > /p/ \quad \text{CUPPAM} > \quad \text{copa} 'cup'
/\textipa{tt}/ > /t/ \quad \text{GUTTAM} > \quad \text{gota} 'drop'
/\textipa{kk}/ > /k/ \quad \text{PECCATUM} > \quad \text{pecado} 'sin'
/\textipa{dd}/ > /d/ \quad \text{ADDUCERE} > \quad \text{aduzir} 'adduce'
/\textipa{bb}/ > /b/ \quad \text{ABBATEM} > \quad \text{abad} 'abbot'

etc.

(B) **Geminate sonorants simplified later** (10th-11th c):

/\textipa{nn}/ > /\textipa{n}, n/ \quad \text{ANNUM} > \quad \text{Sp. año, Ptg. ano} 'year'
/\textipa{ll}/ > /\textipa{l}, l/ \quad \text{BELLUM} > \quad \text{Sp. belo, Ptg. belo} 'pretty'

etc.

- Argued to be due to the step-wise reranking of FAITH with respect to the constraints of the sonority hierarchy. (Latin: FAITH >> Sonority; **Stage A**: *Obs* 
  \mu >> FAITH >> *Son* \mu; **Stage B**: *Obs* \mu, *Son* \mu >> FAITH)

(ii) **Degemination** in Fula (West African) (Sherer §4.2):

Underlying moraic consonants are degeminated after a long vowel

/\textipa{Ni}^\mu w^\mu i^\mu/ \rightarrow [Nii.wi] ([Ni.wi])

- More costly to parse a consonant’s mora than a vowel’s mora, implying *C_\mu >> *V_\mu.

2 **“NO LONG VOWEL”: How should it be formulated?**

In this section I consider three alternatives for the formulation of NLV, shown in (6a,b,c).

Given that *V_\mu prohibits a vowel from bearing a mora, the possibility arises that NLV is the simple double violation of *V_\mu ((6a) in the handout).

(6) **Three options:**

(a) **First option:**

\textbf{NO LONG VOWEL} as double violation of *V_\mu
However, this cannot be if the sonority hierarchy is universal, as appears both desirable and necessary. If NLV were merely the double violation of $*V_{\mu}$, gemination or compensatory lengthening of consonants should never obtain because the addition of a vocalic mora should be preferable to that of one of lesser sonority.

**BUT:** there are cases where gemination interacts with vowel length:

(i) **Gemination** in Wiyot, Koya and Fula (Sherer 1994):
   (All require $*\text{LONG-V} >> *\mu_{\text{cons}}$; Sherer’s terms for my NLV and $*C_{\mu}$)

   (A) **Wiyot** (California) (Sherer §2.7)
   /sopel/ ‘in the middle’ $\rightarrow$ [sop.pel] (*[soo.pel])

   (B) **Koya** (Dravidian) (Sherer §4.3)
   Geminates give rise to underparsing of a vowel mora:
   /o\mu o\mu ndu/ $\rightarrow$ [ot.to(ndu)] (*[oo.to(ndu)])

   (C) **Fula** (West African) (Sherer §4.2)
   /le\mu fi\mu/ $\rightarrow$ [le.f[.i] (*[lee.fi])

   (Compare (4-ii), Fula gemination: total ranking must be NLV >> $*C_{\mu} >> *V_{\mu}$.)

• These cases show that NLV can’t reduce to $*V_{\mu}$, because the double violation of $*V_{\mu}$ should still be optimal when compared to the single violation of $*C_{\mu}$ if $*C_{\mu} >> *V_{\mu}$.

(ii) **Chamorro** (Crosswhite 1996)
   Gemination of initial consonant of a -CV suffix to provide a coda for a stressed syllable (p. 61):
   ‘ettigu’ ‘short’  etti’gonña ‘shorter’
   ‘daŋkulo’ ‘big’  daŋku’lonña ‘bigger’

• Requires some constraint against long vowels to be higher ranked than one against geminate consonants; but, if order of constraints of sonority hierarchy fixed, and NLV is double violation of $*V_{\mu}$, then Chamorro data is unexplained.
(iii) Italian, Sardinian, Late Spoken Latin / proto-Romance have geminate consonants but not long vowels:

<table>
<thead>
<tr>
<th>Italian</th>
<th>Sardinian</th>
</tr>
</thead>
<tbody>
<tr>
<td>affila</td>
<td>‘to sharpen’</td>
</tr>
<tr>
<td>asse</td>
<td>asse</td>
</tr>
<tr>
<td>latte</td>
<td>latte</td>
</tr>
<tr>
<td>martellu</td>
<td>marteddu</td>
</tr>
</tbody>
</table>

Late Spoken Latin / proto-Romance

ANNUM ‘year’
BELLUM   ‘pretty’

- The fact that these languages lack phonologically long vowels may be attributed to the ranking of NLV >> *Cµ, but, again, if sonority hierarchy fixed and *Cµ >> *Vµ, these data lack explanation if NLV is mere double violation of *Vµ.

(6b) Second option:

NO LONG VOWEL as local conjunction of *Vµ: (à la Smolensky 1995)

\{*Vµ \& *Vµ\}

- Should be regarded as a last resort. Conjunction is a powerful theoretical mechanism whose boundaries are under debate (Fukazawa and Miglio 1996, Miglio and Fukazawa 1997).
- Self conjunction is particularly dangerous because in principle it allows for one of the basic tenets of OT to be undermined. Recursive conjunction potentially invalidates the insight that the number of violations of a lower ranked constraint is irrelevant if the candidate satisfies a higher ranked one another does not.

(6c) Third option: (A more cautious formulation, and the version advocated here)

NO LONG VOWEL as vowel associated to two moras:

*Vµµ

- Coexistence of *Vµ and *Vµµ; but theoretically informed motivation for this move.
- Allows for the sonority hierarchy to be maintained, retaining insight that at least vowels will be moraic/syllable peaks in a language.
- Allows gemination or compensatory lengthening to occur at the expense of vowel lengthening (Fula, Wiyot, Koya, Chamorro).
- Allows for an explanation of the phonological inventory of languages with long consonants but not long vowels (Italian, Sardinian, Late Spoken Latin / proto-Romance)
3 Conclusion

- Empirical data supports the claim for the necessity of a constraint militating against long vowels, here termed NO LONG VOWEL.
- Third option, $*V_{\mu\mu}$, most favorable on empirical and theoretical grounds: sonority hierarchy maintained in OT, allows gemination over vowel lengthening, avoids dangers of self conjunction.
- Coexistence of $*V_{\mu}$ and $*V_{\mu\mu}$ required; theoretically informed motivation for NLV.
- $*V_{\mu}$ is part of sonority hierarchy, which tells us what segments may bear a (single) mora; it can’t tell us, however, when a vowel may bear two moras. $*V_{\mu\mu}$ does this.
- $*V_{\mu\mu}$ is similar to $*V_{\mu}$, in that each governs the association of moraic structure to a vowel; however, each is distinct, with $*V_{\mu\mu}$ independently rankable.

This paper has taken a close look at the validity of a constraint that has been employed by many researchers, and has motivated a precise and theoretically informed characterization of NLV that allows for the incorporation and maintaining of the sonority hierarchy in OT. Though the formulation of NLV proposed here introduces an apparent redundancy into the set of constraints $Con$, in that separate but similar constraints operate on a monomoraic and bimoraic vowel, this is a necessary move that allows for an explanatory account of the language data treated here.

References
Smolensky, Paul. 1995. On the internal structure of the constraint component $Con$ of UG.
   Handout of talk at Johns Hopkins University. Spring 1995.