THE TYPOLOGY AND DEVELOPMENT OF TAUTOSYLLABIC SONORANT + OBSTRUENT ONSETS AND RELEVANCE FOR THE SONORITY SEQUENCING PRINCIPLE

DANIEL MANN
CITY UNIVERSITY OF NEW YORK, GRADUATE CENTER
1. Introduction

One of the goals of phonological theory is to explain why certain sound patterns are more common than others. For example, there is a typological asymmetry in the types of consonant clusters allowed in onset of a phonological word. Tautosyllabic onset clusters that have an initial obstruent followed by a glide, liquid, or nasal are fairly common (e.g., English *smell*, Hebrew *dlaat* ‘pumpkin’, Lakota *błę* ‘lake’). However, the reverse is not true, tautosyllabic sonorant (R) + obstruent (O) onsets as in Russian *rta* ‘mouth.gen’ and Klamath *lk’om*- ‘charcoal’ are relatively rare. Their rarity is traditionally assumed to be the result of the Sonority Sequencing Principle (SSP), a proposed phonological universal constraint on syllable organization (Berent et al. 2007; Clements 1990). Syllables generally have low sonority elements, such as obstruents, at the margins and high sonority elements, like vowels, as nuclei:

(1) Sonority Hierarchy (Clements 1990; Parker 2002; Blevins 1995)\(^1\)

<table>
<thead>
<tr>
<th>Vowels</th>
<th>Glides</th>
<th>Liquids</th>
<th>Nasals</th>
<th>Fricatives</th>
<th>Plosives</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a, i, u/</td>
<td>/j, w/</td>
<td>/l, ɹ, r/</td>
<td>/n, m/</td>
<td>/s, f, θ/</td>
<td>/b, t, k/</td>
</tr>
</tbody>
</table>

For instance, English */tɛm.pl/ rises in sonority to its nucleus */æ/ and then falls, ending on the */p/.

A form like */tɛmpl/ would violate the SSP (as a monosyllable) since there is a second rise to the */l/. In order to make this form licit, a vowel must be epenthesized or the */l/ must be syllabic as in English: */tram.pl/\(^2\)

---

\(^1\) Different researchers divide the broad sonority classes up differently. Clements (1990) collapses stops and fricatives in the class of obstruents. Lennertz (2010) keeps them distinct. Parker (2002) mentions that many don’t distinguish glides and vowels since the distinction follows already from syllable structure.

\(^2\) */tɛmpl/ may be syllabified as */tɛmp.ł/ with the */p/ being associated with the previous syllable. However, I assume it adheres to the maximal onset constraint, so */p/ would be associated with the syllabic */l/ (Goslin & Frauenfelder 2001, and references within). Whether the second syllable is */pl/ or */l/ is not
The SSP is claimed to be a phonological universal that is active in the minds of all humans, i.e., part of Universal Grammar (UG; Chomsky 1965; Berent et al. 2007). However, languages like Russian and Klamath show clear violations of the SSP; can the SSP be a true phonological universal? Or is it simply a universal tendency? True universals should be exceptionless and should be active, dynamic parts of every human grammar, while typological tendencies are generalizations, descriptive, and are not part of the grammar (Blevins 2010; Kiparsky 2008). There does seem to be a strong tendency for languages to conform to the SSP, despite the existence of languages with #RO clusters, and the work of Iris Berent may suggest that speakers may have SSP intuitions even without exposure to complex clusters (Berent et al 2007; 2008).

Evolutionary Phonology (EP; Blevins 2004) assumes that typologically common sound patterns are the result of natural sound change, not innate phonological knowledge like that of Universal Grammar. Natural sound changes are based in articulatory, perceptual, and cognitive biases that are not linguistic specific. From this perspective, the typological preference for syllables that adhere to the SSP, seems to be related to the association of a syllable to rises and falls in loudness, i.e., phonetic sonority. This preference, however, is just that: a preference.

This paper attempts to provide evidence that the SSP is not a phonological universal by looking at two prerequisites of true phonological universals: 1) that phonological universals should be exceptionless and 2) that they are actively present in all speakers’ phonological grammar. To account for the first issue, I explore the numerous languages that have #RO clusters and provide evidence that the clusters are, in fact, tautosyllabic and therefore true crucial to the discussion of sonority since both syllabification are consistent with the sonority hierarchy. What would be unexpected from the SSP would be /tae.mpl/.
violations of the SSP. Furthermore, this paper pursues an explanation for the relative rarity of #RO clusters by exploring historical evidence and the sound changes that may have led to them. Looking at the historical evidence it becomes clear that #RO clusters come about via only one historical pathway, that of interconsonantal vowel loss. Furthermore, in the environment of #R_OV, the phonetic properties of sonorants inhibit the loss of a vowel, therefore a #RO sequence is less likely to be phonologized than less “marked” onsets.

To deal with the second prerequisite, that SSP is not part of every speakers synchronic phonological grammar, this paper examines the work of Iris Berent and her colleagues. Her work provides provocative evidence suggesting that the SSP is present in the minds of all speakers, even when their intuitions seem as though they cannot be projected from their language’s lexicon. Despite the intriguing arguments made from the UG perspective, SSP projection effects can be explained by phonetic factors, rather than innate phonological knowledge.

The paper is divided into six sections. Section 1 looks at several languages that are argued to have #RO clusters and assesses whether the clusters are tautosyllabic. In Section 2, I review the relevant historical data (where available) to determine which diachronic patterns give rise to #RO. Section 3, explores the phonetics and phonologization of unstressed vowel loss and explains why this natural sound change rarely gives rise to #RO clusters. In Section 4, I evaluate the SSP’s status as a phonological universal. I pay deserved attention to the work of Iris Berent and colleagues and offer an alternative account for her data. Section 5 presents potential avenues

---

3 Some theories of phonology still argue for phonological universals despite the existence of exceptions. For instance, Optimality Theory (OT; Prince & Smolensky 1993/2004) handles universal violations by integrating competing constraints, some of which force languages to be faithful to the linguistic input (faithfulness constraints) and others that force languages to be comply with the phonological universals (markedness constraints). A purely phonological account, however, makes no predictions as to how or why exceptions come about, or the frequency at which they should be expected.
for further research, specifically, why languages might allow #RO to arise. The final section is the conclusion.

2. #RO Languages

In my typological survey, I collected data from ten languages that have tautosyllabic sonorant + obstruent onset clusters. An onset cluster is defined as a word initial sequence of two or more tautosyllabic consonants without any intervening phonological material. Phonological, unlike phonetic, material must be used in phonemic contrast, syllabification, tone, or some other phonological process. For example, the excrescent vowel in a [C.CV́] sequence in a language that always has initial stress would not be considered phonological. It is more likely a result of gestural mistiming and is unnoticed by speakers (Davidson & Stone 2003).

The starting point of the survey was Greenberg’s 1978 and Kreitman’s 2008 typologies of onset clusters. Though they consider many other languages to have #RO onsets I do not agree with all of their assessments (for example, see section 3.4.2. for a discussion on sesquisyllables). I chose the languages because of their geographic and historical diversity, where “historical diversity” is defined in terms of relatedness (e.g., there is no known genetic/historical connection between Indo-European and Austronesian), as well as diversity in terms of syllable shape, presence of clusters, and lexical base form (e.g., Proto-Indo-European is reconstructed with several consonant clusters while Proto-Maipurean reconstructions have none). The language

---

I will follow the following conventions in this paper:
Where the phonemic/phonetic status was not stated in the data source or where the distinction is not explicitly relevant, I will use italics for all non-English terms. I have standardized all of the forms to a broad IPA. Unless otherwise noted * refers to reconstructed form. Where I use * to mean an illicit form, I note it with a footnote.
I use the following abbreviations for natural classes:
Consonant (C), Vowel (V)
Sonorant (R)= glides (G), liquids (L), and nasals (N)
Obstruents (O)= stops (T/D), fricatives, both sibilant (S) and non-sibilant (F)
families I included in the survey are Austronesian, Indo-European, Kartvelian, Maipurean, Oto-Manguean, and Penutian.

(2) #RO languages

<table>
<thead>
<tr>
<th>Family</th>
<th>Sub-group</th>
<th>Language</th>
<th>#RO Cluster Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austronesian</td>
<td>Formosan</td>
<td>Tsou</td>
<td>ɻtuu ‘loquat tree’</td>
</tr>
<tr>
<td></td>
<td>Central Malayo-Polynesian</td>
<td>Selaru</td>
<td>ltuat ‘small basket for steaming rice’</td>
</tr>
<tr>
<td>Indo European</td>
<td>Slavic</td>
<td>Russian; Polish</td>
<td>rta ‘mouth.gen’; wgate ‘to lie’</td>
</tr>
<tr>
<td>Kartvelian</td>
<td>Karto-Zan</td>
<td>Georgian</td>
<td>mgeli ‘wolf’</td>
</tr>
<tr>
<td>Maipurean (Arawakan)</td>
<td>Southern Maipurean</td>
<td>Piro (Yine)</td>
<td>wkata ‘direction toward this place’</td>
</tr>
<tr>
<td>Oto-Manguean</td>
<td>Mixtecan</td>
<td>Trique</td>
<td>rkweʃaʔ ‘tejamanil (a type of roofing)’</td>
</tr>
<tr>
<td></td>
<td>Zapotecan</td>
<td>Villa Alta Zapotec; Western Highlands Chatino</td>
<td>rbizaʔ ‘hooked stick’ lkaʔ ‘room’</td>
</tr>
<tr>
<td>Penutian</td>
<td>Plateau Penutian</td>
<td>Klamath</td>
<td>lk’om ‘charcoal’</td>
</tr>
</tbody>
</table>

I excluded languages that had only homorganic #NO clusters because it is difficult to determine the phonological representation of clusters that are often orthographically represented as a nasal + obstruent onset. These sequences could be truly NO, or they could be prenasalized obstruents, NO, postploded nasals, NO, or a sequence of a syllabic nasal followed by a non-
tautosyllabic obstruent (Cohn & Riehl 2008). I focused most of my attention on clusters that contain either taps/flaps, trills, rhotic approximants or laterals as the first member of the cluster and stops or fricatives as the second member. There is less risk of analyzing a liquid + obstruent sequence as anything other than LO. There are no rhoticized or lateralized obstruents so, unlike NO clusters, it is unlikely that LO clusters would be analyzed as unary elements (there may be disagreement with regard to syllabicity of the sonorant but I will discuss it on a language specific basis below).

2.1. Data Sources

The following publications were the principle sources for the language data:

5 Downing (2005) and Ladefoged & Maddieson (1998) argue that there is no phonetic difference between prenasalized consonants and nasal + stop sequences (though there may be a phonological difference). In constrast, Cohn & Riehl (2008) contend that there are differences in the durational ratios between prenasalization, postplosion, and true consonant clusters.

6 I was cautious of clusters with sibilants, due to their phonological and phonetic exceptionality (see Fleishhacker 2001 for a detailed discussion). However, I have yet to come across a language that has #RS clusters but no #RT or #RF clusters, so until the data forces me to, I will not be treating sibilants in C2 positions as special.

7 When a different source is used, or when it is important to single out a particular source, the data will be cited in the text.
(3) Data sources

<table>
<thead>
<tr>
<th>Language</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsou</td>
<td>Blust (ongoing), Tsuchida (1976), Tung (1964), Chen (2002)</td>
</tr>
<tr>
<td>Russian &amp; Polish</td>
<td>Krísto &amp; Scheer (2005), Bethin (1998), Sussex &amp; Cubberley (2006), Russian National Corpus (Ongoing)</td>
</tr>
<tr>
<td>Villa Alta Zapotec</td>
<td>Yatzachi variety (Butler 1980; Butler 2000), Zoogocho variety (Long &amp; Cruz 1999), Betaza variety (Olivares 2009)</td>
</tr>
</tbody>
</table>

2.2. Implicational facts

All of the sampled languages adhere to Greenberg’s (1978) implicational statement that a language with #RO will have #OO, and a language with #OO will have #OR.\(^8\) If a language has a sonorant + obstruent cluster we can assume it will also allow sonority plateaus, a tendency other typological surveys of onsets have found (Kreitman 2008; Morelli 1999). The implicational statement is consistent with an EP account. As I show in Section 3, all of these clusters types arise through the same process of interconsonantal vowel loss.\(^9\) In phonetic terms, #RO is the least likely onset to arise in this environment (see Section 3.4). A sound change that leads to #RO necessarily leads to #OO.

---

\(^8\) Trique is slightly complicated, though. Native consonant clusters in Trique are more restrictive than the other languages in the survey. The only native onset clusters that adhere to the SSP are gj-, sn- and sw-. The rest are either plateaus with an initial s- (or falls if one divides obstruents into stops and fricatives) or falls with an initial r- (DiCanio 2008).

\(^9\) I don’t claim that #OO and #OR come about exclusively by interconsonantal vowel loss, only that they can and do arise through syncope and that vowel loss that leads to #RO also affects the less typologically marked clusters.
The #RO clusters do not provide any evidence of systematic gaps. #RO clusters are not only rare cross-linguistically but are rare language internally, therefore, it is premature to claim that there are true gaps and/or tendencies. Many #RO clusters that do not exist are likely historical accidents. Russian, for instance, lacks velarized lateral + coronal onsets. There is no evidence that this reflects any systematic restriction in its phonological grammar (unlike English speakers perception of tl as either tr or tel; Pitt 1998). Experimental evidence suggests that Russian speakers, in terms of syllabification, perceive the unattested consonant clusters the same as the attested clusters (Berent et al. 2007). Considering that Russian clusters that violate the SSP account for only 1.06% of all clusters with liquids, some gaps are to be expected (Proctor 2009). Other gaps seem to be just as accidental, Piro lacks #LO clusters, but l and r are rare in word
initial positions in Piro (Matteson 1965). Similarly, $p$ as a second consonant is quite rare in the Oto-Manguean languages, which is not surprising since Proto-Oto-Manguean is reconstructed as having no bilabial obstruents to begin with (Rensch 1966).

2.3. Tautosyllabicity

As I am arguing these languages are exceptions to the Sonority Sequencing Principle it is necessary to demonstrate that the clusters in question are, in fact, tautosyllabic. A phonological SSP has no way of accounting for tautosyllabic clusters. If the SSP is an active component of all formal grammars, speakers should repair an illicit #RO cluster, with one option being syllabic sonorants. If, however, the sonorant consonants are treated as syllabic by speakers then the SSP has not been violated.

2.3.1. Austronesian

2.3.1.1. Tsou

Historically, Austronesian languages prefer disyllabic lexical bases (Blust 2007). Modern Tsou seems to adhere to this preference as 90% of Tsou’s lexical bases in Blust’s comparative dictionary are disyllabic. The rhotics in the following forms, if syllabic, would counter Tsou’s disyllabic trend:

---

10 The liquids occur in a word initial position in Spanish loans words but $l$ never occurs word initially in native words. $r$- is a 3rd person masculine prefix so inflected words do begin with the rhotic. However, the pronominal prefix, replaces the initial root segment ($h > r$) rather than “attaching” itself, it doesn’t give $rO$ an opportunity to arise: himka ‘to sleep’; nimka ‘I sleep’; pimka ‘you sleep’; rimka ‘he sleeps’.

11 The synchronic grammars allow them as polysyllabic Spanish words undergo vowel loss when adopted into the languages, e.g., Spn. refajo > Zapotec (Betaza) rpâg ‘a type of skirt’ (Olivares 2009).
(5) Tsou disyllabic/bimoraic\textsuperscript{12} #RO clusters

a. \textit{qvora} ‘flying squirrel’
b. \textit{qse\textalpha} ‘ears’
c. \textit{mvoe} ‘dry in the sun’

Canonical Tsou syllables are composed of a vocalic nucleus and an optional onset of one or two (and no more) consonants. Codas are not permitted. If the sonorants are assumed to be strictly consonantal, special syllable rules are not required nor are there any exceptional syllable shapes present. If, however, sonorants are considered syllabic the phonology is complicated by the presence of a class of nuclei that cannot take an onset (or perhaps even stranger, nuclei that \textit{must} be followed by another nucleus e.g., \textit{fr.o.he} ‘moon’).\textsuperscript{13}

Tsou syllable reduplication also indicates that these #RO clusters are true consonant clusters. The reduplication targets the initial syllable and suffixes it to the base:

(6) Tsou syllable reduplication: \#OCVV > \#OCVVOCV

a. \textit{fŋuu} > \textit{fŋu-fŋu} ‘big head’
b. \textit{smuu} > \textit{smuu-smu} ‘much due’
c. \textit{skuu} > \textit{sku-sklu} ‘necklaces’

#RO clusters in (7) are treated the same as the #OR and the #OO clusters above in (6):

\textsuperscript{12} Tung (1964) considers sequences of identical vowels to be heterosyllabic based on stress assignment, reduplication, and affixation facts. Wright and Ladefoged (1994) analyze the vowels in terms of weight. Under this analysis, Tsou would have bimoraic roots rather than disyllabic roots. The relevant fact is that Tsou has lexical bases that either have two syllables or two morae. Under a moraic analysis, the rhotic behaves like other onsets in that is not moraic.

\textsuperscript{13} Bell (1970) states that syllabic consonants are typically more restrictive in terms of syllable complexity than vocalic nuclei in that they typically do not permit complex onsets or codas. However, the inability for sonorants to take even a simple onset in Tsou and the complete lack of patterning with the uncontroversial nuclei makes any claim of syllabicity suspect.
(7) Tsou syllable reduplication: #ROVV > #ROVVROV\textsuperscript{14}

\begin{itemize}
  \item a. mciúu > mciúu-mcůu \#mcúu-ml \#mcúu-mu ‘ears of crops’
  \item b. mcóó > mcóó-mcůo \#mcóo-ml \#mcóo-mo ‘big eye’
  \item c. ŋsóó > ŋsóó-ŋso \#ŋsóo-ŋl \#ŋsóo-ŋo ‘much spring’\textsuperscript{15}
\end{itemize}

As there is no difference in how the clusters are treated there seems to be little justification for treating the sonorants as syllabic. A #ROV base, where the sonorant receives stress, would be strong evidence for the presence of syllabic consonants in Tsou since the language has penultimate stress. Stress would have to fall on the sonorant to make the form more consistent with Tsou’s disyllabic preference. I have yet to encounter such a word.

\subsection*{2.3.1.2. Selaru}

Selaru is less studied than its Austronesian cousin, Tsou, and thus the evidence is less clear. Coward (2005) and Coward & Coward (2000) have the most data on the phonology and grammar of the language and they claim syllabicity only for vowels, yet never explicitly deal with the status of syllabic sonorants.

Root size is much more variable in Selaru than in Tsou; there are more monosyllabic and trisyllabic forms in the former. Selaru similarly has penultimate stress so forms like rsus ‘3rd pl.suck’ and rfíl ‘3rd pl.choose’ could be expected to have stress on the sonorant if it were to be syllabic. However, it seems as though stress is assigned before morphological processing occurs as the difference between ʔamána ‘octopus’ and ʔáma-na ‘father.his’ suggest.

Many of the Selaru rhoic #RO clusters arose through the addition of the third person plural affix \textit{r}. The lack non-vocalic stress in rsus and rfíl could be due to either the lack of

\footnotesize{\textsuperscript{14} \# in (7) indicates an unattested/illegal form.  
\textsuperscript{15} Chen’s (2002) data focuses on a dialect of Tsou where the rhotic has been lost. Since no dialect of Tsou has a lateral, its only sonorants are nasals.}
syllabicity or to pre-morphological stress assignment. However, *lkusy* [lkusǐ] ‘urn’ has no morphological breaks and does not have stress on the sonorant, suggesting the sonorants are not syllabic.

Perhaps the best evidence for treating the sonorants as exclusively consonantal comes from the interaction of morphology with root clusters. With the exception of glides, Selaru maximally allows two consonants in the onset. When *r-* is affixed to a root that already has a maximal consonant cluster a vowel is epenthesized:

(8) Selaru #CCC epenthesis: /R-CCV/ > [RVCCV], /R-CV/ > [RCV]

a. /r-knam/ > [raknam] ‘3rd pl.eat’
a’. /r-kuty/ > [rkutǐ] ‘3rd pl.bite’
b. /r-brai/ > [rabrai] ‘3rd pl.refuse’
b’. /r-bai/ > [rbaǐ] ‘3rd pl.go’
c. /r-mtaut/ > [ramtaut] ‘3rd pl.scare’
c’. /r-ma/ > [rma] ‘3rd pl.come’

If the *r-* were capable of syllabicity, the vowel’s insertion would be completely unmotivated. The vowel is only inserted in a #CCC sequence to conform to Selaru’s syllable shape inventory. The obstruent *r-* which is a 1st person plural verbal prefix, patterns similarly:

(9) Selaru #CCC epenthesis: /O-CCV/ > [OVCCV], /O-CV/ > [OCV]

a. /t-bren/ > [tabren] ‘1st pl.play’
a’ /t-ba/ > [tba] ‘1st pl.go’
b. /t-knam/ > [taknam] ‘1st pl.eat’
b’. /t-karia/ > [tkaria] ‘1st pl.work’
2.3.2. Indo-European

#RO clusters in Russian and Polish are widely accepted as being tautosyllabic, based on experimental evidence, poetry, and comparisons with Slavic languages that clearly have syllabic sonorants (Berent et al. 2007; Davidson 2011; Kreitman 2008).

Perhaps some of the clearest evidence for #RO tautosyllabicity in forms like Rus. lgat‘ to lie’ and mgla ‘haze’ is the experimental work of Berent et al. (2007). While the purpose of the experiment was to show that English speakers had SSP intuitions even without information from the lexicon (I will discuss Berent’s work in further detail in Section 4.1.), Russian speakers provided clear-cut syllabic judgments for nonce words like lba. Overwhelmingly, Russian speakers perceived lba as monosyllabic while perceiving its epenthetic counterpart, leba as disyllabic. This pattern held for clusters with rhotics and nasals as well.

Unlike many of the other languages in the data set, the Slavic languages have well-documented and easily accessible poetry. Poetry is useful in making syllabification judgments as many poets organize the linear and prosodic structure of poems based on syllable count (also known as meter).

(10) Syllabification of Russian poem

lgut-se-kra-na-ko-var-nu-je-poz-ne-ru
u-t jit-ziz-ni-nas-ks/u-sha-sob-tfak
ras-poz-na-li-be-du-slij-kom-pozd-no-mu
f[lo-3e-de-lat-nam-zit-dal-je-kak

Лгут с экрана коварные познеры,
Учит жизни нас Ксюша Собчак.
Распознали беду слишком поздно мы.
Что же делать нам, жить дальше как? (Степашиной, В.
“Обращение к Русским Мужчинам 23-го Февраля”)
The poem in (10) (only partially reproduced here), alternates lines of eleven and nine syllables. A syllabic lateral would render the first line in (10) as having twelve syllables, which would conflict with the rest of the poem.

Finally, when Russian and Polish are compared to Slavic languages that have syllabic sonorant consonants, there is a clear difference. Czech, Serbian, Macedonian, and Slovenian all allow monosyllabic words with the nuclear, and definitionally syllabic, sonorants. Conversely, the Russian and Polish cognates always exhibit a vowel associated with the sonorant:

(11) R ~ VR alternation in Slavic languages (Proctor 2009)

<table>
<thead>
<tr>
<th>Syllabic R language</th>
<th>Polish</th>
<th>Russian</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Serbian <em>pɾst</em></td>
<td>palets</td>
<td><em>pɨrst</em></td>
<td>‘finger’</td>
</tr>
<tr>
<td>b. Czech <em>vɨk</em></td>
<td>wilk</td>
<td><em>volk</em></td>
<td>‘wolf’</td>
</tr>
<tr>
<td>c. Macedonian <em>hɾb</em></td>
<td><em>gərb</em></td>
<td><em>gorb</em></td>
<td>‘hump’</td>
</tr>
<tr>
<td>d. Slovenian <em>sɾp</em></td>
<td><em>sɨerń</em></td>
<td><em>sɨerp</em></td>
<td>‘sickle’</td>
</tr>
</tbody>
</table>

Furthermore, in both Polish and Czech stress is predictable, Polish stress is always penultimate and Czech’s is always initial. Consider the following cognate pair: Czech *tɾvat* and Polish *trwáte* ‘last’. The Polish form carries stress on the vowel. If the sonorant were syllabic, as in Czech, it should be associated to the rhotic.

2.3.3. Kartvelian

In much of the Georgian literature, the sonorants in (SSP reversal) consonant clusters are often called syllabic. Butskhrikidze (2002) does so and appeals to presence of a vowel in other Kartvelian languages and synchronic syncopation (word internal vowel deletion) data to suggest that the syllabicity of the vowel transfers to the sonorant: *mercxal > mercxl-is* ‘swallow.gen’, cf. *k’amat > k’amat-is* ‘debate.gen’. However, it seems that many of those that use this term are
suggesting some sense of *phonetic* syllabicity, rather than phonological syllabicity. Cherci (1999) states that the sonorants “sound syllabic” but also claims that they don’t count as nuclei for meter purposes. The evidence Butskhrikidze provides does not show that Georgian speakers treat the sonorant as syllabic, only that they treat sonorants and obstruents differently with respect to vowel syncope. She seems to be arguing that the realization of the sonorant may be more acoustically vocalic or sonorant, not that it is acting as a phonological syllable nuclei (Ritter 2006).

According to Nepvue (1994) and Robins & Waterson (1952) Georgian speakers never perceive the acoustic peaks of sonorant consonants as syllable peaks, native speaker syllable intuition only treats vowels as syllable nuclei. Since disyllabic words receive initial stress, the lack of stress on the sonorant in *rtwa* ‘to spin’, *rva* ‘eight’, and *mze* ‘sun’ matches with the reported speaker intuition (Robins & Waterson 1952; Butskhrikidze 2002). Articulatory evidence also suggests that Georgian consonant clusters should be considered true clusters and not syllabic sonorant + obstruent sequences. Goldstein et al. (2007) found C-center effects (where the timing lag between the final consonant and the nucleus gets shorter as consonants are added to a cluster) for Georgian clusters with sonorants. This lag was not found with Tashlhiyt Berber where any consonant may act as a phonological syllable nucleus.

### 2.3.4. Maipurean

Like Georgian, Piro sonorants have been called syllabic. Matteson (1965) claims there is a syllabic allophone, not just for sonorants, but for all consonants. While there may be an

---

16 Since the acoustic distinction between syllabic and consonantal sonorants has yet to be established it is unclear what is meant by “sounds” syllabic. The work of Pouplier & Beňuš (2011) demonstrate the difficulty in showing a clear acoustic difference between syllabic and non-syllabic sonorant consonants. See Goldstein et al. (2007) for a different approach for an approach which integrates articulatory timing in clusters.
excrecent vowel in these clusters there is little evidence that suggests they are treated as phonological vowels, and therefore phonological syllable peaks, by Piro speakers. Lin (1998) claims they never operate in any phonological processes or rules like word stress or phrase level rhythmic rules.

The pitch contour Matteson (1965) provides in her phonological description indicate that segments in these complex clusters are not phonologically syllabic. Falls and rises in pitch are always associated with a full, phonological vowel, never a consonant:

(12) Piro tone contours

a. 2 1 2 1 2 1
   txako nato prani
b. 1 1 2 2 1 1 2 2
   kin-fikale, mfiir-fikale

Furthermore, stress is never placed on a consonant, sonorant or otherwise. Piro has primary stress on the penultimate syllable with secondary stress on the initial syllable. Tertiary stress occurs on every other syllable starting with the initial syllable (and moving toward the end of the stress group). If the stress group has an odd number of syllables there will be two unstressed syllables before the penult. The stress group tá.palú.jka.ná.wa.thí.ma.ná.nu.mtá.na.tná.ka ‘it is said that his canoe was going alone was going along again’, only complies with the aforementioned stress rules if vowels are syllable nuclei (Lin 1998).

In (13b), below, all clusters are broken up into syllabic consonants (since Piro does not allow codas the consonant cannot be associated with the previous syllable). The result would be a stress group of eighteen syllables, but we would expect the syllable wa to receive stress and na to be unstressed, contra the produced form. Similarly, in (13c) only the sonorants in SSP
violating clusters are considered syllabic, which again makes the wrong predictions. \( m \) would directly receive stress leaving \( ta \) unstressed.

(13) Predicted stress patterns of syllabic consonants

\[
\begin{align*}
a. & \quad \text{tá.pá.lú.ʃ.jka.ná.wa.thí.má.ná.nu.mtá.ná.na.tná.ká} \\
& \quad \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \\
b. & \quad \text{tá.pá.lú.ʃ.jka.ná.wa.t.hí.má.ná.nu.mtá.ná.na.tná.ká} \\
& \quad \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \\
c. & \quad \text{tá.pá.lú.ʃ.jka.ná.wa.thí.má.ná.nu.mtá.ná.na.tná.ká} \\
& \quad \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma} \overset{\ddot{\sigma}}{\sigma}
\end{align*}
\]

Similarly, words like \( \text{nsó} \sim *\text{ń.so} \) ‘genipa’ and \( \text{wpowráulu} \sim *\text{úpowráulu} \) have exceptional stress patterns if sonorant consonants are syllabic.

2.3.5. Oto-Manguean

2.3.5.1. Trique

Ignoring the sonorant consonants under consideration, Trique syllable shape is (C)(C)(C)V(H) where H is a laryngeal. Much like Tsou, positing a syllabic sonorant in a word like \( \text{rkwe}^3 \text{tfa}^2 \text{ś} \) ‘temajanil (a type of roofing)’ or \( \text{si}^3 \text{rku}^3 \text{tśi}^3 \) ‘purple’ seems unwarranted. In doing so, there would be a nucleus that can take neither an onset (since anything that precedes the sonorant would be either a vowel, laryngeal, or a word boundary) nor a coda (since only \( h \) and \( ? \) are allowed in the coda and sonorants in Trique clusters are followed by non-laryngeal consonants). Furthermore, sonorants are never found in a word final position, unlike the uncontroversial nuclei.

If we assumed a syllabic sonorant, a word like \( \text{jah}^{32} \text{rma}^3 \text{hśi}^3 \) ‘sweet’ would be four syllables; no word in DiCanio’s (2008) dissertation is larger than 3 syllables. All of this is in
addition to the fact that the sonorants are never stressed. Trique has word final stress so the combination of no word final sonorants and requirement of word final stress means that only vowels will receive stress, so there at least seems to be little phonological need for syllabic sonorants.

If sonorants could bear tone, then there may be evidence for their syllabicity but there is little evidence that they do so. If an #ROV sequence had a rising tone (e.g., #ROV\(^{34}\)) there would be clear evidence of syllabicity because the only monosyllables that allow rising tone have coda h. A rising tone on a #ROV sequence is an indication that the word in disyllabic (DiCanio 2008). However, tone rises are never found and even tone falls are infrequent with #RVO sequences:

(14) Trique tone and #RO clusters

a. \textit{rkwiit}^5 ‘grasshopper’
b. \textit{rka}^4 \textit{le}^3\(^3\) ‘mayor’
c. \textit{rkwe}^3 \textit{ruj}^2 ‘back of the knee’
d. \textit{si}^3 \textit{rku}^1 \textit{fih}^3 ‘purple’

In the ATWG dictionary and DiCanio’s (2008) dissertation, most of the #ROV sequences bear only one tone.\(^{17}\)

\textbf{2.3.5.2. Highland Chatino}

Chatino has similar coda restrictions and stress placement to Trique so if consonants could act as nuclei we would need to assume that they are exceptional in that they cannot take onsets, codas,

\(^{17}\) There are some #ROVV sequences that bear two tones such as \textit{rkwa}a\(^{32}\) ‘colander’, but the vowel is always long (or heavy). Since even simple onsets + long vowels have tone rises and falls this is not a problem: \textit{jo}^3^2 ‘sugarcane’; \textit{ri}^5^2 ‘harvest’.

There is one #CROV sequence, \textit{xrto}\(^{32}\) ‘plantain’, that has a falling tone. Since falling tone is ambiguous it would not be evidence of syllabic sonorants. \textit{rto}^2^3 would be, but there are no cases of this.
or be word final. Unfortunately, the tonal evidence is not as clear. Any level tone or sequence of tone that is allowed on simple or SSP adhering clusters is also allowed on SSP reversal clusters:

(15) Highland Chatino tone

a. \textit{lkwi}^{45} `tepache (a type of liquor)`
   a' \textit{da}^{45} `question marker`

b. \textit{lt}a^{21} `armful`
   b'. \textit{?}a^{21} `a lot`

c. \textit{lk}a^{23} `be`
   c'. \textit{br}a^{23} `time`

d. \textit{lt}i^{2} `narrow`
   d'. \textit{ku}^{2} `to eat`
2.3.6. Penutian

Barker (1963) remarks that nasals and liquids do not count for stress pattern. Yet, he refers to them as syllabic because they occur with pitch. However, from his discussion on Klamath tone, it is not clear that the sonorants are, in fact, syllabic. For the most part, they carry the same tonal level as an adjacent vowel. The only variation in this is when the sonorant follows a “juncture” (e.g., a word boundary). In this case, they always have tone [2]. The lack of pitch variation on the sonorants does not clarify syllabicity. It’s possible that the lower tone of [2] could just be the default tone for initial prosodic boundaries. If pitch variation or contours could be shown for sonorants, in the same way there is on the vowels, there may be evidence for syllabicity, but this evidence is lacking.

According to Blevins (1995) all Klamath syllables must have onsets, so if the sonorants are syllabic they would be exceptions to this pattern. The lack of stress and the restrictiveness that is seen with tone does not seem to give us any convincing reason to introduce complexity into the phonology of Klamath syllable structure.

2.4. Summary

While #RO clusters are not common, there are several languages where there is a solid case for true SSP violating clusters. It’s unclear how these languages can be accounted for under a pure UG account. A phonological SSP should require that all of these languages syllabify the sonorant, yet none of these languages do so. Russian and Polish speakers eschew this correction even though their mother language and other Slavic languages have syllabic sonorants. Most of the South Slavic languages #RO clusters exhibit a syllabic sonorant where Russian and Polish have a cluster: Slovenian *rt, rt*tf* ~ Russian *rta, rt*if*e ‘mouth’. Furthermore, a phonological SSP
makes no predictions about how these clusters arise, at what frequency they will be present, or what conditions might cause them to develop.

3. The evolution of #ROV

In order to understand why #RO clusters are typologically rare, we should understand the historical pathways that led to the development of these clusters. A UG account does not account for the exceptions to proposed universals, and so we have no understanding of how exceptions may develop. Evolutionary Phonology, on the other hand, argues that typologically uncommon sound patterns are learned just as easily as common patterns, but natural sound change will lead to certain patterns becoming more common than others (Blevins 2004). The distribution of #RO clusters has the feel of natural sound change since the onsets are rare both language internally and cross-linguistically, meaning the factors that led them to be rare across all languages led them to be rare within languages. In fact, despite all of the potential pathways that could be possible, the only attested pathway is that of interconsonantal vowel loss.

3.1. Language data

3.1.1. Austronesian

Proto-Austronesian (PAN) is reconstructed as having mostly disyllabic lexical bases of CVCV(C). PAN had no consonant clusters in its inventory, unlike other languages in the survey, so the clusters arose without internal analogy. Throughout the various daughter languages unstressed vowels were lost, creating a diversity of cluster types, including #RO:
In addition to clusters that arose in root words, the Austronesian languages also had affixes lose unstressed vowels, resulting in heteromorphemic consonant clusters. The loss of the vowel in the active voice affix ma- in Tsou and the 3rd person plural affix ra- (from Proto-Malayo-Polynesian *siDa ‘they, them’; Coward 2005) in Selaru led to #RO clusters:

(18) Prefix vowel loss

a. *ma + *puSaN > Tsou m-pusku ‘twenty’

b. Proto-Malayo-Polynesian (PMP) *siDa + *bayaD > Selaru r-bayar ‘to pay’

c. PMP *siDa + *piliq > Selaru r-fil ‘to choose’

d. PMP *siDa + *kawit > Selaru r-kait ‘to drag’

e. PMP *siDa + *susu > Selaru r-sus ‘to suck’

3.1.2. Indo European

Slavic #RO clusters come from what is commonly known as “the fall of the jers.” The jers were short high vowels that were lost when they did not receive primary stress (see the alterations in (19f) and (19i) below; also see Proctor 2009; Yearley 1995; Bethin 1998; and

18 Unless otherwise stated, proto forms are Proto-Austronesian.
19 Though the precise qualities of the jers are not completely known, the vowels, orthographically ъ and ъ, respectively, are often described as short high vowels (Yearley 1995). They may have also been lax, /ʌ/ and /u/. They have their roots in the Indo-European *i and *u. (Bethin 1998).
references within for a more detailed discussion on *jer* loss. Like Austronesian vowel loss, the loss of the *jers* created several new consonant clusters in both Polish and Russian:

(19) Proto-Slavic *jer* loss

  a. *sūna* > Rus. sna ‘sleep.gen’
  b. *dīne* > Rus./Pol. dnja ‘day.gen’
  c. *pīsa* > Pol. psa ‘dog.gen’
  d. *kūto* > Rus./Pol. kto ‘who’
  e. *lūgati* > Rus. lga’t; Pol. wga’t ‘lie.inf’
  f. *lūba* > Rus. lba; Pol. wba ‘forehead.gen’ (cf. lūbū > lob ‘forehead.gen’)
  g. *rūvati* Rus. rvat ‘tear, rip’
  h. *rtūta* > Rus. rta ‘mouth.gen-sg’ (cf. *rūtū > rot ‘mouth.nom’)
  i. *rūtūtī* > Rus. rtul; Pol. rtā ‘mercury’
  j. *mūgla* > Rus. mgla; Pol. mgwa ‘haze’

Unlike Austronesian, however, consonant clusters were already widespread in both Proto-Slavic and its mother language, Proto-Indo-European (PIE). PIE #OR clusters (e.g., *kʷrei- ‘buy’) were widespread and the language even permitted #OO (e.g., *pter- ‘wing’) and #RR (e.g., *mregju- ‘short’) clusters. Several PIE consonant clusters have persisted, with minimal changes, into the modern Slavic languages:

(20) Proto-Indo-European clusters > Modern Slavic

  a. PIE *bʰre₂t̚er > PSl *bratr̚ja > Rus./Pol. brat ‘brother’
  b. PIE *sneigʷ-o- > Rus. sneg; Pol. enieg ‘snow’
  c. PIE *klőwā > Rus. slava ‘fame’

20 The loss also sometimes resulted in consonant palatalization. Russian has maintain a phonemic contrast of velarized (“hard” /r/, /l/, /m/, /n/) and palatalized (“soft” /ɾ/, /ɫ/, /m̥/, /n̥/) variants of consonants.

Draft
3.1.3. Kartvelian

The historical data on Georgian is not as clear as the Austronesian or Slavic data, yet for many #RO clusters there is a reconstructed #ROV sequence, suggesting Georgian underwent the same process of interconsonantal vowel loss as well:

(21) Common Kartvelian (CK)/Georgian-Zan (GZ) > Georgian

a. CK *lag- > rgav ‘plant’
b. GZ *lab- > lboba ‘soaking until soft’
c. CK *rekw- > rka ‘remark’
d. CK *lexw- > lxoba- ‘to thaw, to melt’
e. CK *rekw- > rk’uma ‘to say’
f. CK *reyw- > ryveva ‘demolish, collapse, to pour, pull down’

However, because consonant clusters are so widespread through the Kartvelian languages the proto forms for many modern #RO clusters are reconstructed without an intervening vowel:

(22) *#RO > #RO

a. GZ *rt- > rtva- ‘join’
b. GZ *rk’en > rk’en- ‘to fight, wrestle’
c. CK *(s)ʒe- > rdze ‘milk’
d. CK*mʒe- > mze ‘sun’

Forms like those in (22a-d) lead Butskhrikidze (2002) to believe that Proto-Kartvelain had syllabic sonorants and the modern forms are derived from syllabic sonorant consonant + obstruent sequences. Vogt (1958), however, contends that the clusters are the result of vowel loss. It is possible that both are correct. As shown in (21), some clusters did arise from the loss of a vowel and Bell (1978) argues that all syllabic sonorants are the result of vowel loss. So, if true, syllabic sonorants could have been a stage in Georgian #RO cluster development.
Further evidence for a historical vowel comes from bases with multiple reconstructions. *rts'eva* ‘to shake’ and *rkvam* ‘to cover, be covered’ have both #RVO and #RO proto forms: GZ *refʰ-x-/*rts'-* and CK *rekw-/*rkw-am-*, respectively. Finally, the earliest branching Kartvelian language, Svan, has a vowel where CK has an #RO sequence: Svan *ložë* ‘milk’; *mɔz/miʒ* ‘sun’, cf. (22c,d).

3.1.4. Maipurean

Payne (1991) reconstructed Proto-Maipurean to have a syllable shape of (C)V(C) with only n and h occurring in the coda. This is in stark contrast to Piro which allows almost any combination of segments and allows sonority plateaus and sonority reversals of #NO and #GO clusters (Matteson 1965). Piro is the only Maipurean language that deviates (at least so radically) from the CV syllable shape (Aikhenvald 1999). So, like Austronesian, Piro shows evidence of clusters developing from the loss of an interconsonantal vowel de novo:

(23) PMaipurean > Piro

a. *kahitʰi > ksiri ‘moon’
b. *kiri > čri ‘nose’
c. *mata > mta ‘skin’
d. *nene > nne ‘tongue’
e. *kafa > kfiwna ‘armadillo’
f. *teni > tni ‘breast (milk)’
g. *kasiukʰi > kfiyoçri ‘cayman’

---

21 Maipurean lost vowels at word edges, as well. The proto form for *mko* ‘cloud’ is *amī-nV-ko-ri*. It seems as thought the loss of the initial vowel did not lead to a tautosyllabic cluster, however. Proto-Asháninka, a closely related Southern Maipurean language, is reconstructed as having the form **menko-ri**, so Proto-Piro probably had no initial vowel in this word either. In theory, the loss of word initial vowels could have given rise to #RO clusters but it seems less common since #VC is a rarer syllable type and a cluster can only arise in a #VCVC sequence if multiple vowels are lost.
3.1.5. Oto-Manguean

Rensch (1966) reconstructs #NO clusters and palatal + consonant (YC) clusters, though the exact specifications are unknown. Otherwise the language is constructed to have a syllable structure of CV(?). The Zapotecan branch, which includes Chatino and Zapotec, is reconstructed with a syllable shape of CV(?), without the #YC and #NO onsets (Fernandez de Miranda 1995; Suarez 1973; Swadesh 1947). It is generally well-accepted that Highland Chatino, Trique, and Villa Alta Zapotecan consonant clusters all came about by the loss of unstressed interconsonantal vowels (Campbell 2013; DiCanio 2008; Swadesh 1947):22

(24) Proto-Zapotecan vowel loss

a. *Lagaʔ > Chatino lkaʔ23 ‘leaf’
b. *yaga > Chatino yka2 ‘tree’
c. *yuʒi > Chatino ysiin4 ‘sand’
d. *luʒu > Chatino wsin21 ‘beard’
e. *Loba > Chatino lwaa45 ‘sweep’
f. *lasiʔ > Chatino lti2 ‘skinny’
g. *luʒeʔ > Chatino lte2 ‘tongue’
h. *Laʔana > Yatzachi Zapotec lton ‘hunger’

Oto-Manguaean languages range from conservative varieties which maintain the Proto CV structure and allow polysyllabic forms to innovative varieties that have drifted toward monosyllabicity and allow few bases larger than two syllables:

(25) Zapotecan #RO cluster varieties ~ non-#RO cluster varieties

a. Yatzachi Zap lban ~Isthmus Zapotec libana ‘sermon’
b. Yat.Z lbax ~Sierra de Juárez Zapotec libixxi ‘behind’
c. Yateé Zap. wʒidyʔ ~ Santo Domingo gʊʒiʔdʒán ‘I will laugh’
d. Yateé Zap. bʒidyʔ ~ Santo Domingo bʊʒiʔdʒán ‘I laughed’

22 The reconstruction has been more detailed at the level of the Zapotecan branch. Fernandez de Miranda’s used Chatino, but not V.A. Zapotec, as one of the bases for reconstruction, so examples of V > ə / C.CV are more abundant for PZ> Chatino.
Like the Austronesian languages, several clusters #RO clusters arose through an inflectional prefix that lost a vowel, such as Proto-Zapotecan habitual marker *ri- > r- (Campbell 2011).

3.1.6. Penutian

A genetic relationship between Klamath and other languages along the North American West Coast has yet to be widely accepted among most linguists (DeLancey & Golla 1997). This makes any claim to the origin of #RO clusters in Klamath slightly weaker than that of Austronesian or Slavic. However, there has been extensive work done to connect the languages of the Southwest Canadian/Northwest US Coast that has been fruitful (Delancey et al 1988; Tarpent 1997). The available data demonstrates the same pattern that is seen with the aforementioned languages; #RO alternates with #RVO:

(26) Klamath ~ Proposed Penutian cognates

<table>
<thead>
<tr>
<th>Klamath</th>
<th>Sahaptian</th>
<th>Nez Perce</th>
<th>Proto-Yokuts</th>
<th>Proto-Tsimshianic</th>
<th>Nez Perce</th>
</tr>
</thead>
<tbody>
<tr>
<td>lk’om</td>
<td>lāk’im</td>
<td>mīta-at</td>
<td>*nek’il</td>
<td>*lēXw</td>
<td>wē:wtuk</td>
</tr>
<tr>
<td>ndan</td>
<td>‘coals, charcoal’</td>
<td>‘soot’</td>
<td>‘three’</td>
<td>‘run, move quickly’</td>
<td>‘to sleep’</td>
</tr>
<tr>
<td>ṫkena</td>
<td>‘spear’</td>
<td>keetis</td>
<td>*nek’il</td>
<td>*lēXw</td>
<td>wē:wtuk</td>
</tr>
<tr>
<td>pk’isísap</td>
<td>‘mother’</td>
<td>píke</td>
<td>‘(his) mother’</td>
<td>‘underside’</td>
<td>‘to camp overnight’</td>
</tr>
<tr>
<td>ptisap</td>
<td>‘father’</td>
<td>píst</td>
<td>‘(his) father’</td>
<td>‘underside’</td>
<td>‘to camp overnight’</td>
</tr>
<tr>
<td>wle</td>
<td>‘run’</td>
<td>wílé</td>
<td>*nek’il</td>
<td>*lēXw</td>
<td>wē:wtuk</td>
</tr>
<tr>
<td>nk’ey</td>
<td>‘bullet’</td>
<td>*nek’il</td>
<td>*lēXw</td>
<td>‘to sleep’</td>
<td>‘to camp overnight’</td>
</tr>
<tr>
<td>l’qh-</td>
<td>‘down to the ground’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
understanding how the phonetics might be a limiting factors in the development of #RO should be less complicated and more straightforward.

3.2. Unstressed vowel loss

Unstressed vowel loss has a clear phonetic base. Unstressed vowels are shorter than their stressed counterparts which makes them better candidates eventual loss. In a coupling model of syllable structure (Browman & Goldstein 1988; Nam et al. 2009) the gestures for the consonant and vowel in a CV sequence are initiated synchronously which means the consonantal gesture will always partially overlap with the vocalic gesture. The shorter a vowel is the more “hidden”, or less perceptible, it will be (Chitoran & Iskarous 2008).

(27)

\[\begin{array}{c}
\text{a.} \\
\text{b.}
\end{array}\]

In both (27a) and (27b) the CV gestures are synchronous. The long vowel has less of its overall duration obscured by the consonantal gesture. The short vowel will be difficult to perceive as a result of being almost completely overlapped by the initial consonant.

Furthermore, a short vowel requires movement of the tongue over a shorter period of time, which leads to “undershoot”, where the vowel does not reach its acoustic target (Lindblom 2005).

3.3. Phonologization of clusters as tautosyllabic

The phonetic weakening of the vocalic gesture and a continuum of vocalic realizations along a hyper-hypoarticulated scale is the perfect environment for a change in the phonological grammar (Blevins 2004). With inputs of [CiC…], [CəC…], [C’C…], [CC…] the child, in creating their phonological system, may decide that there is no intervening vowel and assume /CC…/. If this
happens, any presence of a vowel in hyperarticulated speech may be assumed to be an allophonic variant of the underlying CC structure, the opposite of the previous generations’ interpretation of a vowel being deleted in the output (or on the surface form); /CVC…/ > [CC…]. This process would be an instantiation of CHOICE in EP. The weakened vowel can be interpreted several ways. It could be interpreted as the presence of a vowel, consonantal release, or a gestural mistiming that allows for a brief period of an unobstructed vocal tract. The phonologization process could also be CHANGE. As I will discuss below in Section 3.4.1 and Section 4.1. R̆VO sequences are likely to be perceptually ambiguous with #RO sequences. A listener may hear #R̆VO as #RO (though, as I will discuss, this is also an inhibiting factor to their developing).

Of course, vowel reduction does not require the loss to be phonologized. English, for instance, has vowel reduction and words like potato are often realized as [pʰtʰ]ayto. Orthography, prescriptive norms, among other factors may prevent a cluster interpretation (Blevins 2004).

3.4. Why are #RO clusters rare?
Vowel loss is a relatively common process, so why are #RO clusters rare? Vowel loss will not give rise to the same ambiguity in all environments, we should expect that some C̆VC sequences will be more prone to a CC interpretation than others. One factor is the effect of voicing. Devoiced vowels are even less perceptible between consonants, consequently, C̆VC should be an optimal environment for vowel loss (Chitoran & Iskarous 2008). Sonorants are spontaneously voiced (Chomsky & Halle 1968), therefore are not expected to devoice the following vowel. In the process of vowel loss, we would expect vowels to be more resistant to phonological loss in a #[+Voice]VO environment.
As I will discuss in the following section, unlike obstruents, sonorants have vocalic properties associated with them; they are complex segments that have both consonantal and vocalic gestures, which are often be asynchronous (particularly with liquids; Sproat & Fujimura 1993 Pouplier & Beňuš 2011). This “vocalic element” could be a lack of oral constriction, a period of “relatively stable-state formant structure” (Ladefoged & Maddison 1998), and/or a dorsal gesture (Sproat and Fujimura 1993). Essentially, #RO should always be able to be interpreted as #RVO by a naïve learner/listener. I also consider the possibility that vowel reduction being phonologized as either sesquisyllables or as syllabic sonorant consonants reduces the typological frequency of #RO.

3.4.1. Vocalic properties of sonorants

In the languages learning process, a child should always have the choice of interpreting a #RC sequence as an #RVC sequence because of the vocalic elements inherent to sonorants. The rhotic trill [r] has a series occlusions are separated by brief periods of “vocalic intervals” (Jaworski & Gillian 2011). These open phases have clear formant structure similar to that of vowels (Ladefoged & Maddison 1998):
Similarly, with the tap [ɾ] when the tongue is moving toward and away from the alveolar ridge (or teeth) there is, what Savu (2012) refers to as, a “vocoid”. This vocoid, or vocalic element, is where the is little to no constriction of the vocal tract, similar to the [r] in (28), so in a spectrogram you can see energy concentrated in formant regions.

The lateral gesture is made of two asynchronous components: one apical and one dorsal (Sproat and Fujimura 1993). The dorsal gesture, often associated with vowels, shifts toward the syllable nucleus (i.e., “gestural affinity”) while the apical consonantal gesture is attracted to the margins, a phenomenon that occurs with all of the sonorant (multi-gesture) segments (Byrd et al. 2010). When there is no adjacent vowel, the presence of the dorsal gesture could be equated to

---

23Female Russian speaker. Data from forvo.com, a website where users can access and/or upload pronunciations of lexical items. The accuracy and nativeness of the pronunciation was confirmed by other native speakers: [http://www.forvo.com/word/%D1%80%D0%B4%D0%B5%D1%82%D1%8C/#ru](http://www.forvo.com/word/%D1%80%D0%B4%D0%B5%D1%82%D1%8C/#ru)

24Krackow 1999 found that the constriction (consonantal) gesture precedes a vocalic gesture for liquids (a dorsal gesture for [ɾ] and pharyngeal gesture for [ɬ]) and is preceded or synchronous with the nasal vocalic gesture (the lowering of the velum) when the sonorant precedes a vowel. When it follows a vowel, the
that of a vowel. Especially, since segments at the edges of prosodic boundaries tend to be lengthened, the lateral in an absolute onset position will be realized as longer than the word internal variant (all things being equal, of course; Oxley et al. 2006). The longer a lateral segment is, the more likely it is to be confusable with a vowel since vowels are long and kinematically slow segments. The liquids may never completely lose the reduced vowel, and if they do, there are gestures inherent to the segments that have more of a chance of being interpreted as vowels by future generations. Though they may be shorter in duration, glides would be similar to the approximant liquids in being composed of a consonantal gesture and a vocalic gesture.

Nasals also have clear formant structure associated with them, though perhaps not to the same extent as liquids. They are composed of consonantal gesture that completely blocks oral airflow, but they have a secondary velic gesture that allows air to flow through the nasal passage (Byrd et al. 2010). This gives the nasals continuous, non-turbulent airflow and so they have fairly clear formant frequencies (Ladefoged & Maddieson 1998). The further back the articulation the less salient the antiresonance effects are, which enhances the confusability of a nasal stop and a nasal vowel (Ohala & Ohala 1993).

Because there is always a vocalic gesture associated with the sonorants, but not necessarily with obstruents, we should expect that languages that undergo unstressed vowel loss might target #O’O > #OO for phonologization, and not #R’O > #RO. Unlike Russian and gestures switch places with the dorsal, pharyngeal, or velic gesture coming before the consonantal gesture.

25 The lateral is often also confused for a glide. For instance, in Polish, #lO > #wO.

26 With voiced #OO sequences there is often a release burst associated with the initial consonant. Depending on the timing of the articulators there may be a period where formant bands can be seen. Kreitman (2008; for Hebrew) and Chitoran (1998; for Georgian) refer to these as “vowel-like” and “transitional”. Unlike with sonorants where the vocalic elements (formant structure, periodicity, open
Polish, some Slavic languages, such as Bulgarian, the complete phonologization #RO clusters did not occur:

(29) Russian ~ Bulgarian

a. rva ~ rov ‘ditch’
b. rʒi ~ rʒen ‘rye’
c. lgat ~ lʒa ‘to lie’
d. mstij ~ mʒi ‘revenge’

Bulgarian did, however, phonologize the vowel loss between obstruents: PSl *dūva > dve ‘two’, *pūtika > ptiča ‘bird’, and *kišo > fiš ‘what’.

Even when the vowel had previously been lost, there is still a chance that the sonorant will later be interpreted as two distinct segments. As an Eastern Slavic language, Belarusian completely lost the jers in an #RVO environment. However, subsequent generations interpreted #RO with a prothetic vowel, i.e., not in its historical position:

(30) Russian ~ Belarusian

a. ɨba ~ ilba ‘forehead.gen’
b. rtui ~ irtus ‘mercury’
c. mgla ~ imhla ‘haze’

3.4.2. Other results of vowel loss

Because #RO clusters come about through the loss of a vowel (i.e., a syllable nucleus) the phonological syllabicity of the vowel could be transferred to the sonorant, resulting in a syllabic sonorant. This process seems to have occurred in Germanic languages and Southern Slavic

---

vocal tract, etc.) are inherent to the segment, the transitional vowel is not always present. See the appendix for Kreitman’s (2008) spectrograms of #OO and #O'O.
languages (Scheer 2009; Bell 1978). In fact, Bell (1978) claims that all syllabic consonants come about through the loss of a vowel.

The reduced vowel may also become sesquisyllabic. Sesquisyllables are neither disyllabic or monosyllabic but, rather, are composed of a minor syllable and a major syllable. The canonical shape of a sesquisyllabic word is CəCVC, with the minor (or pre-) syllable having more restrictions on syllable shape (e.g., no codas), a reduced phonemic inventory (often only one phonemic vowel is allowed), and lack of stress (Bennett 1995). Many Tibeto-Burman languages, such as Balti, and Austroasiatic languages, such as Khmer and Khasi are traditionally argued to have sesquisyllables, contra Kreitman (2008) and Greenberg’s (1978) treatment of them as #RO clusters (Matisoff 1990). Austronesian languages are reconstructed with sesquisyllables, indicating a historical stability (Michaud 2012). Matisoff (1990) argues some of the Tibeto-Burman languages have undergone a cycle of syllable contraction and expansion: Pre-Proto-Tibeto-Burman *sya-myak > Proto-Tibeto-Burman *sɔmyak > *smyak > Lahu mêʔ > mêʔ-fi > (‘hypothetical future’) mafī ‘eye’. Whether or not the language is stable or goes through cycles of decay and rebirth, sesquisyllables are a potential option for a language that undergoes vowel reduction, limiting the frequency of #RO clusters.27

27 A recent dissertation on sesquisyllables by Butler (2014) disputes a sesquisyllabic analysis for some of the Southeastern Asian languages, including Khmer. She argues that there has been a wide variety of definitions for sesquisyllables, many of which may be too inclusive. She contends that some of the languages, like Khmer, can be analyzed as having consonant clusters with an excrescent vowel. She bases her argument on the fact that Khmer vowels in sesquisyllables do not reach their acoustic target and are heavily influenced by adjacent consonants. If Khmer is not a sesquisyllabic language it would be a language with sonority reversals, necessitating inclusion in this list. Further analysis of the phonological facts and of her phonetic argument will be necessary.
3.5. Summary

The rarity of these clusters seems to be the result of a confluence of factors. First of all, there needs to be vowel loss. While unstressed vowel loss is common, a language that develops #RO clusters must also have a stress system that allows stress to not be on an initial syllable. Any language that has fixed primary stress or fixed penultimate stress with numerous disyllabic forms would not be expected to develop #RO clusters. The reconstructed stress systems in the survey indicate bear out these predictions: Proto-Oto-Manguean stress is on the ultimate syllable (Rensch 1966), Proto-Slavic stress is free and mobile (Sussex & Cubberley 2006), and Proto-Austronesian has two root types: roots with penultimate stress and roots with ultimate stress (Wolff 1993).

Additionally, that vowel loss must be phonologized. Sonorants have properties that inhibit this phonologization, so not every language that has vowel loss will phonologize #RO, like the Bulgarian example. Finally, the language cannot take an alternate pathway, like that of syllabic sonorants or sesquisyllables.

4. The SSP: a phonological universal or universal tendency?

#RO clusters in all of these languages clearly violate the SSP, bringing its status as a true universal into question. It also seems as though we can account for the rarity of the clusters by appealing to the phonetic-based sound change, negating the need for a phonological SSP.

However, data from Iris Berent et al. (Berent et al 2007; Berent et al. 2008; Gómez et al. 2014) indicates that speakers treat consonant clusters differently depending on their sonority profile, even when their language allows only sonority rises. Berent et al. argue that this is an indication that a phonologically universal SSP is present in all speakers and an active component
of all human grammar. While the effects are real, it is premature to suggest that the results can only be explained by appealing to a universal, synchronic preference for syllable types. Rather, there may be an explanation for the data based on perception and perceptual confusability.

4.1. Berent and the SSP

English doesn’t have phonological small sonority rises (bnif), sonority plateaus (bdif), or sonority reversals (lbif), yet English speaking subjects do not treat all clusters as equally “bad”. Instead, they perceive #ON clusters as monosyllabic more often than #OO and #RO is almost never perceived as monosyllabic. Furthermore, English speaking subjects confuse #RO with #RVO at a higher rate than #OO ~ #OVO (and much more than #ON ~ OVN). Korean speakers, whose language only has #OG clusters have similar intuitions. Their intuitions adhere to not just the languages’ phonotactics, but also, the SSP:

28 Unless sibilant + stop is considered a plateau. However, the phonetics and phonology of sibilants are exceptional and are often the only “plateaus” that a language allows (Fleischhacker 2002). For this reason, and others, many leave sibilants out of the equations when discussing SSP and SSP intuitions.
(31) Korean and English speakers and consonant clusters

The graph is taken from Berent et al. (2008) and covers the experiments with Korean speakers and English speakers. The white shapes are an identity test where subjects judged whether stimuli of the form lebif~lbif were the same or different. The black shapes represent a syllable task that had subjects give their intuitions on whether words like lbif were monosyllabic or disyllabic. Falls were of the sonority shape lbif; plateaus, bdif; small rises, bnif; and large rises blif. The relative rankings of preferred to dispreferred syllable onsets would be as follows: #OG/#OL > #ON > #OO > #RO.

Berent et al. (2007) argues phonetic factors cannot explain the data. For instance, lbif is confused for lebif but the reverse is never true, lebif is never perceived as lbif, which, they claim, should be expected if perceptual confusion is the source of the projection effects. However, bi-directionality is not necessary for perceptual confusion. lbif should be confused for lebif because of the inherent phonetic properties mentioned in Section 3.4.1. In forms lebif, elbow [ˈɛlbəʊ] or blow [bləʊ], the dorsal vocalic gesture shifts toward the vowel and can therefore be associated with the e (gestural affinity; Sproat & Fujimori 1993). Contra Berent et al.’s claim, lebif
shouldn’t be confused with *lbif* and the perception of [ɛlbow] and [below], from inputs of [ɛlbow] and [blow], are not necessarily expected from a phonetic account. Conversely, a form such as *lbif* that is comprised of a sonorant surrounded by silence (word onset silence and silence of the stop gap)\(^\text{29}\) is predicted to have a default interpretation of RV (or VR) due to the multi-gesture phonetic composition of the sonorant (e.g., the vocalic gestures plus apical gestures of *r*, *l*, and *r*; oral closure of nasals, lip rounding of *w*, etc.).

These vocalic elements are not present (or at least not required) in stop + stop clusters. Thus, English and Korean speakers may not be (just) perceptually epenthesizing a vowel, Berent et al.’s claim, but rather, they are interpreting the already present vocalic elements as a reduced vowel. The fact that Russian speakers do not perceive any vocoid is due to their language specific grammar, i.e., they have learned to ignore the vocalic elements.

Berent and colleagues also found that the disyllabic forms of the unmarked clusters such as *benif* are confused for the monosyllabic forms more often than the disyllabic forms of the marked clusters. *benif* is confused for *bnif* more often than *bedif* is confused for *bdif*, a result they believe further supports a phonological account, especially since the schwa length was the same for both tokens. Again, this is not inconsistent with a phonetic explanation. While it is possible that the vowel in #OɛO could be interpreted as the release of the initial O or as a gestural mistiming, a reduced vowel is often longer than either of those vocoids.\(^\text{30}\) Kreitman (2008), found that, for Hebrew, the epenthetic vocal elements were never longer than 30ms, shorter than any of reduced vowels in her data. *benif* may be perceived as monosyllabic because the transition between a sonorant and an adjacent vowel is not sharply discontinuous and because of the co-articulatory effects between the vowel and sonorant. For instance, the velum may be

\(^{29}\) Or reduced airflow with fricatives.

\(^{30}\) And unlike sonorants, where it is often difficult to tell where a sonorant ends and vowel begins, the closure and burst of stops give a clearer “beginning” and “end”.
lowered in anticipation of the nasal gesture, creating a nasalized vowel/vocoid (Ohala & Ohala 1993). In an OəRV environment, with the obstruent partially obscuring the reduced vowel, the speaker may misperceive the beginning of the consonantal gesture and assume the reduced vowel is, in fact, part of the sonorant.

4.2. Predictions of a phonetic account

A phonetic account based on the degree of vocal tract aperture (or nasal airflow) makes similar predictions to the SSP, in that #OR is preferred to #OO and both are preferred to #RO. However, with respect to sibilants, voiced obstruents, and rhotic taps there should be some slight variations.

Sibilants are often ignored in the sonority literature, particularly because they are somewhat exceptional both typologically and phonetically. For many of the world’s languages if there is only one marked onset it will involve a sibilant (Fleishhacker 2001). For instance, English does not allow #OO onsets where the first member is a non-sibilant, like *[tkæl], *[fpray] or *[θtaɹ]32, yet allows [skæl], [spray], and [staɹ]. Phonetically, they are loud segments, yet have a small vocal tract opening (Ladefoged & Maddieson 1998). Since they are clearly fricatives, they should pattern with the obstruents, yet typologically this is not the case with respect to the SSP. Fleishhacker (2001) argues that their particular phonetics helps explain their typological distribution. Their loudness allows them to be perceptible in most environments so they can be found in a wider range of positions throughout the phonological word. However, they are not resonant like a vowel, therefore, they don’t carry as well, making them weak as syllable nuclei (Henke et al. 2012). A phonetic account of sonority affords a special status for

31 Liquids altered the F2 and F3 of surrounding vowels (Ohala & Ohala 1993). All of these coarticulatory effects could effect the lack of a vowel percept.
32 * here indicates illicit form.
sibilants; it’s unclear how a purely phonological account deals with sibilants beyond ignoring them or offering post hoc analyses.

With a phonetic account there should also be a difference between voiced and voiceless stops in word initial position. As mentioned in Section 4.1. and by Kreitman (2008), there may be a brief period of an unobstructed vocal tract when a speaker is transitioning from one stop to another. In a #TT cluster (or perhaps even #TD, where T is a voiceless stop and D is voiced) voicing should not be present during this brief period. Without voicing, there is much less of a chance of the percept of a vowel. In a #DD sequence, however, the onset of voicing before the initial stop remains through both gestures, enhancing the likelihood of a #DVD interpretation. A phonetic account predicts that #TT is less likely than #DD to be perceived as disyllabic and confused for #OVO.\textsuperscript{33}

Finally, the phonetics of a second consonant rhotic flap (Cr) should be expected to be interpreted differently than other sonorants. The transition from an initial consonant to a lateral, rhotic approximant, nasal, or glide is continuous and shows no abrupt stoppage of air. The flap, however, has a very brief, but complete, closure of the vocal tract. As seen in the spectrograms in (32a) and (32b), the flap has a more noticeable transition that could aide in the percept of a reduced vowel. We should expect that speakers who have no complex clusters might be more likely to perceive Cr as CVr than Cl as CVL. A phonological SSP would not predict a difference between the two.

\textsuperscript{33} Many linguists do not divide the class of obstruents or stops into voiced vs. voiceless. Those that do divide stops consider voiced to be more sonorous than voiceless (Parker 2002). Berent et al (2007) does not distinguish them. From their perspective, both clusters plateau so a UG account should predict them to be the same.
(32) Spanish Cr and Cl

a. [bre...]³⁴

b. [ble...]

³⁴ Both utterances were taken from the same native Spanish (from Spain) speaker on Forvo.com:
http://www.forvo.com/word/brezo/
http://www.forvo.com/word/joaqu%C3%ADn_blake_y_joyes/
5. Summary and questions for future research

A Universal Grammar approach, which claims a universal phonological SSP seems unwarranted. If the SSP is an active, component of all humans’ phonological systems it’s unclear how #RO clusters would arise and why they wouldn’t be interpreted as syllabic. A diachronic approach, couched in Evolutionary Phonology, predicts that they should be rare because of phonetic biases. Cross-linguistically, all #RO clusters came about through interconsonantal vowel loss. Because of the phonetics of vowel loss and the intrinsic phonetic properties of sonorant consonant segments (e.g., their confusability with sonorant + vowel sequences) we should expect that these clusters would be less likely to arise than the typologically unmarked sequences. Yet, work remains to be done. Many of the claims made here still need strong experimental evidence, such as the predictions in Section 4.2. It also remains to be seen if there are any factors that may increases the likelihood of the development of #RO. In the next section, I discuss potential factors.

5.1. Why do #RO languages exist?

#RO is rare because of the phonetic properties of inter-silence sonorants. But what allows a child to disregard the vocalic elements and assume the sonorant is a unary segment rather than a sonorant vowel (or VR) sequence? Structural analogy (Blevins 2009) could play a role in the phonologization of #RO. Blevins (2004) defines structural analogy as:

(33) Structural analogy

In the course of language acquisition, the existence of a phonological contrast between A and B will result in more instances of sound change involving shifts of ambiguous elements to A or B than if no contrast between A and B existed.
An ambiguous #R'O sequence could be interpreted as #RO if there is a regular sound change of \( V > \emptyset/ #C_C\bar{V} \). The learner may give less importance to the fact that there are no preexisting #RO clusters and give greater weight to unstressed vowel loss.

\[(34) \text{ CHOICE} \]

Phonetic realization of
\[
/#RVOV/ > [R'O], [RVO], [RO], [R\bar{a}O]
\]

Historical, register, grammatical alternation
\[
V > \emptyset/ #O_OV; #O_R\bar{V}^{35}
\]

Synchronic grammar
\[
/#RVO/ *#RO
\]

\[
/#RO/
\]

\[
/#RVO/
\]

(34) illustrates the subconscious choices that a child must make in creating its grammar. A language that undergoes vowel reduction in certain contexts (e.g., relaxed speech) may have several options available to them. If the synchronic grammar of the previous generation allows vowel loss with #OO and #RO clusters but does not with #RO clusters, the child could “choose” to analogize the ambiguous #RXO (where X is any relatively unobstructed vocal tract with formant structure) sequences to that of #OR and #OO or to #RVO. However, the factors that might push a child toward one interpretation over the other needs to be explored. The overall frequency of reduced vowels (e.g., a fixed final stress system with unstressed vowel reduction

---

35 Lezgian, a Northwest Caucasian language is currently undergoing a process of vowel reduction leading to onset consonant clusters. Vowels are being lost only after voiceless consonants, a process expected from natural sound change. In light of the phonetic facts and Lezgian, it could be assumed that vowels after voiceless obstruent initial consonants will be the first segments to lost and phonologized as CC, all things being equal (Chitaran & Iskarous 2008).
should lead to an abundance of [C'CV] and [CCV] sequences), frequency or preexistence of consonant clusters, and/or any number of “unnatural” factors (register variation, contact, prescriptive norms) could all play a role.

Further in-depth research also needs to be done to determine what factors may explain the #RVO > /#RO/; /#/O/; /#/əO/ variations. What factors might lead to the phonologization of /#RO/ over a syllabic sonorant consonant or sesquisyllable? One possibility is “attractor” effects. Proto-Oto-Manguean and Proto-Austronesian were predominantly disyllabic (Campbell 2013, Blust 2007). If the vast majority of lexical bases are only disyllabic then ambiguous sequences, that may have arisen through the addition of suffixes or clitics, might be attracted to a disyllabic interpretation (Blevins 2009). Syllabic sonorants consonants or sesquisyllables would not be an option since the driving force for the vowel loss is maintenance of syllable restrictions on bases. The language learning child is faced with a choice; adhere to the syllable preference of words or attenuate to the phonetic cues.

6. Conclusion

#RO clusters are typologically much rarer than #OR and #OO clusters, often held to be a result of the SSP. While the SSP has long been considered a phonological universal that explains phonotactic patterns and syllable shape (Parker 2012), there are numerous problems with the proposal and the field remains divided on its relevance (Parker 2012). Working within the framework of Evolutionary Phonology (Blevins 2004; 2014), I have demonstrated that numerous languages do have clusters that violate the SSP, facts that cannot be explained by a UG account. Furthermore, I have shown that there are similarities in how #RO clusters developed. #RO clusters arose through the loss of an interconsonantal vowel which helps explain the typological
rarity. The lack of multiple pathways and the phonetic facts of sonorants in an #RO environment serve to limit the development of #RO cluster languages.

I also explored the work of Berent and colleagues in order to assess their claims of an active phonological SSP. Berent et al. (2007) found that speakers prefer clusters that adhere to the SSP over those that do not, even when none of the relevant clusters exist in their language. I demonstrated that their results can be explained by perceptual and articulatory factors and that we need not rely on universal phonological knowledge.
References


Appendix

Krietman’s (2008) spectrograms for Hebrew speakers:

**Spectrogram 5.12**: Speaker 1 – *dgálim* ‘flags’.

Spectrogram 5.12 shows the release burst of the *d*. There is no clear formant structure and relatively little energy for the release burst.
For speaker 3, there is a transitional vowel. The formant bands can be seen and there is higher acoustic energy.

Spectrogram 5.19: Speaker – 3 – *dgali* ‘flags’.