Syllable Structure in Umatilla Sahaptin*

Tymoteusz Alan Chrzanowski

A thesis submitted in partial fulfillment of the requirements for
the degree of Bachelor of Arts in Linguistics

Swarthmore College
December 2018

Abstract

Umatilla Sahaptin is a Shaptian language spoken in North-Eastern Oregon by the
Umatilla tribe. While it has some documentation, such as a dictionary and sketch
grammar, there are still acknowledged gaps in the literature. The purpose of this
thesis is to fill one of those gaps with a description of syllable structure. To do this
I compare data gathered from speakers of Umatilla with published accounts of sylla­
ble structure in the mutually intelligible sister language Yakima. After establishing
what is meant by syllable structure and what is known about Yakima, I discuss the
data I gathered. Using word list and sentence elicitations, stories, poems, and songs
I conclude that Umatilla syllable structure is very similar to Yakima's, with some
differences, namely in maximal margins, minimal words, and licensing of sonority
sequence defying clusters.

*I would like to thank my consultants, Thomas Morningowl and Mildred Quaemptes, since it was
their time and knowledge that made this work possible at all. I would also like to thank the
Confederated Tribes of the Umatilla, for the privilege of working with the tribe, specifically the
Language Program, the Education Commission, and the Board of Trustees. In particular I would
like to thank Modesta Minthorn, for her help and guidance through the process of securing per­
mission for my work and scheduling time to do it. I would also like to thank the Eugene M. Lang
Summer Research Fellowship committee, for their generous financial gift which allowed me to
conduct this research. Finally, I would like to thank my adviser, Professor Emily Gasser. Without
her help from the very inception of this project it would never have happened.
Table of Contents

1 Introduction .............................................. 3

2 Background Information .......................... 3
   2.1 Area and Family .................................. 3
   2.2 Syllable Structure ................................ 4
   2.3 Yakima ............................................ 8
   2.4 Umatilla .......................................... 11

3 Research ................................................. 12
   3.1 Data Gathered .................................... 12
   3.2 Field Experience ................................... 12

4 Analysis .................................................. 14
   4.1 Understanding the Data ........................ 14
   4.2 Comparison with Yakima ......................... 15

5 Conclusion ............................................... 17
   5.1 Further Research ................................. 18

6 Appendix .................................................. 20
   6.1 Elicitation Data ................................... 20
   6.2 Songs ............................................... 22
   6.3 Poems .............................................. 24
1 Introduction

Umatilla Sahaptin, also sometimes called Tamalúut or Ichishkfin (hereafter Umatilla), is a Sahaptian language spoken in North-Eastern Oregon by the Umatilla tribe, with between 25 and 125 speakers (Lewis et al. 2015). Given this critically endangered position, there is a strong sense of urgency to document and analyze this language before its unique linguistic insights are lost. Fortunately, there is scholarship that has been done on it by Rude (2014), in the form of a sketch grammar and a dictionary. There is also an active language revitalization program under the Confederated Tribes of the Umatilla (hereafter, CTUIR) government, which has produced enough pedagogical materials to have classes at both the elementary and high school level.

However, there are still several acknowledged gaps in the literature, and it is imperative to fill as many of them as possible while access to native speakers is still available. The goal of this thesis is to give a description of Umatilla syllable structure, first by reviewing related languages and what is already known about Umatilla and then preceding to use data gathered from fieldwork with native speakers that used hypotheses generated from this earlier work to guide it. I conducted this fieldwork in the summer of 2018, and have found that Umatilla largely overlaps with Yakima in terms of syllable structure, but with some differences regarding syllable size and adherence to the sonority hierarchy.

2 Background Information

2.1 Area and Family

The Umatilla Reservation is approximately 271 square miles in the northern Blue Mountains of Oregon, with a population of around 3000 people. Historically, the Umatilla had extensive relationships with other tribes in the area, including the Walla walla and the Cayuse, with whom they would form the Confederated Tribes of the Umatilla in 1864 (CTUIR 2018).

The Sahaptian language family is a part of the Plateau Penutian group\(^1\) in the Pacific Northwest (specifically Washington, Oregon, and Idaho). It has two main branches, Nez Percé and Sahaptin, the latter of which can be further divided into

\(^1\)It has been speculated that Plateau Penutian is related to Uto-Aztecan, but the evidence is not conclusive (Rude 2000)
Northern and Southern Sahaptin. Umatilla is a variety of Southern Sahaptin, as is the mutually intelligible Yakima (Jansen 2010).

![Figure 1: The Umatilla Reservation](image)

Figure 1: The Umatilla Reservation

Plateau Penutian

Sahaptian

Sahaptin Nez Percé

Southern Sahaptin Northern Sahaptin

Umatilla Yakima

Figure 2: Umatilla Family Tree

Much like other languages of the Pacific Northwest, Sahaptin is characterized by polysynthesis and complex verbal morphology, as well as a large consonant inventory and a relatively small vowel inventory (Ibid).

There has been work done on the family since 1862, with Pandosy's description of Pshwanwapam (Ibid). The primary scholarship has been done on Nez Percé (primarily by Noel Rude), Yakima (primarily by Sharon Hargus and Virginia Beavert), and Umatilla. There has been more relevant research done on Yakima than on Umatilla, therefore in section 2.3 I review some of the work on Yakima and how it can be used to inform work done on Umatilla.

2.2 Syllable Structure

The syllable is a widely accepted abstract phonological unit, useful for analysis of a language’s prosodic features like rhythm and stress, as well as explaining cross-
linguistic patterns in segment sequencing (Zec 2003). It also gives motivation and explanation for phonological processes, such as repair strategies like epenthesis. Furthermore, the syllable is an entrenched method of understanding sound patterns for things such as poetic meter.

The components of a syllable are its nucleus, usually a high sonority phoneme (generally a vowel), and the margins, the onset on the left and the coda on the right. The nucleus and coda are grouped together in the rime, giving the following structure:

\[
\text{Syllable} \rightarrow \text{Onset} \rightarrow \text{Rime} \rightarrow \text{Nucleus} \rightarrow \text{Coda}
\]

Figure 3: Syllable Structure (Kessler and Treiman 1997)

When describing syllable structure, it is conventional to use \(V\) for nuclei, regardless of whether it is or must be an actual vowel, and \(C\) for all consonants in margins; thus, \(CV\) can stand in for any syllable with an onset, but not a coda. It is also possible for the onset and/or coda to contain multiple segments, producing complex onsets and codas, so a syllable with a complex onset could be represented as \(CCV\). Any group of adjacent consonants can be described as a consonant cluster, and as well as existing within syllables (as in the case of complex margins), they may span syllable boundaries. So in the word \[matʃ.boks\], the bolded segments are a consonant cluster. Finally, if a segment is optional, this can be shown with parentheses. For instance, if it were possible to have two coda consonants, but only one was required, it would be like this: \(CVC(C)\).

2.2.1 Minimality and Syllable Shape

What constitutes the minimal allowable syllable in a particular language is known as minimality. There is an important distinction to be made when discussing minimality, between the minimal word and the minimal syllable. For instance, as discussed later, Yakima is word minimally biconsonantal, but the minimal syllable is \(V\).

The minimal possible syllable is just \(V\) and many languages, English for one, it is also the minimal syllable allowed. However, some languages require more complex syllables, for instance, Senufo requires an onset (Laughren 1973). Notably,
there are almost no attested cases of languages that specifically require codas. Al­
most universally languages allow onsets, but not vice versa; many languages, such
as the earlier mentioned Senufo, ban codas.

There is at least one notable exception to this pattern however. Arrernte is a
central Australian language that appears to have only codas, and furthermore, re­
quire them (Breen and Pensalfini 1999). While it is technically possible to produce
an account of syllabification in Arrernte with onsets, it is much more complicated
than simply admitting the possibility of an onsetless language2 (Ibid). Breen and
Pensalfini’s theory as to why Arrernte has such an unusual structure is that its abun­
dance of prenasalized stops shifts the phonetically salient edge of the segment from
right to left. They hypothesize that the reason for onsets generally being unmarked
is that most consonants’ acoustically salient edge is their right, making the tran­
sition from onset to nucleus acoustically easier to parse; therefore, if, in the case
of Arrernte prenasalized stops, the salient edge is instead the left, that distinction
moves to the transition from nucleus to coda. This demonstrates that there are
limitations to our theories of syllable structure and we must be open to revising
it, albeit only with good cause. It should also be noted, that this analysis of Ar­
rernte is controversial, but regardless of which analysis one favors, it presents an
interesting case.

The patterns of minimality and maximality provide acceptable syllable shapes,
but phonotactics determines what phonemes can occupy the positions in those
shapes.

2.2.2 Phonotactics

Languages also have rules and restrictions about which of their phonemes can
appear where. This is called phonotactics, and it can be broadly separated into
two categories, rules about position and rules about environment.

Rules about position are directly relevant to syllable structure, in that they
are based on positions within the syllable (or word). For instance, it is intuitive to
English speakers that the syllable [ŋI] is not valid, while [Iŋ] is quite common: this
is because English does not allow engma in onsets, but does allow it in codas. This,
of course varies between languages, and for example [ŋa] is the Vietnamese word
for ‘Russia’. These sorts of rules can also refer to positions within clusters, or even

2It should be noted that Arrernte does not, at first glance, seem to lack onsets, since it has words
that are consonant initial. However, closer analysis indicates that these words originally possessed
an initial vowel, which was then deleted phonologically
a combination of both of these. The word [round] has the cluster [ndl], which, if moved into an onset position, becomes invalid, despite both sounds being allowed in isolation in an onset and in a coda cluster (cf. the language Ndebele, which in its name contains this cluster in an onset).

Rules about environment are about what phonemes can appear adjacent to each other. Take the cluster [kn], while this is orthographically allowable in English onsets, these two sounds are not permitted to be in this order in this position. It is not because of their position in the onset, both are perfectly able to be there (e.g. [kIr] and [nIr]), nor can it be their particular place in the cluster, since we see them in those respective positions in the words [krip] and [snol.kl]; it is not positional, therefore it must be because of the phonemes themselves. However, we do see this cluster across the syllable boundary, in the word [æk.ni] ‘acne’. Of course, it is important to generalize and not just have a list of specific prohibited combinations, so here we might posit that English does not allow Stop + Nasal clusters in onsets. A strong motivation and explanation for many such rules lies in sonority.

2.2.3 Sonority

Sonority describes the amount of acoustic energy in a sound, and it is directly related to how open the vocal tract when producing the sound: the more open the vocal tract, the more sonorous the phoneme. Sonority is relevant to syllable structure because it is a factor in how syllables are perceived. Speakers identify syllables with the number of sonority peaks in a word, thus [pli] is likely to be heard as one syllable, [lpi] might be misperceived as two (Harrington and Cox 2009).

![Graphical Representation of Sonority Peaks](Harrington and Cox 2009)

Figure 4: Graphical Representation of Sonority Peaks (Harrington and Cox 2009)

For this reason, syllables tend to follow a gradient in sonority, so as one approaches the nucleus, the sounds tend to be more sonorous, and less so as one approaches the margins, this is known as the Sonority Sequencing Principle. For
every language there can be generated a sonority hierarchy of all of its phonemes, but generally they conform, more or less, to this pattern:

Vowels > Glides > Liquids > Nasals > Voiced Fricatives > Voiced Stops
> Voiceless Fricatives > Voiceless Stops

Figure 5: General Sonority Hierarchy from High to Low (Zec 2003)

Of course, this hierarchy is not inviolable, two phonemes of the same sonority level can form a sonority plateau (e.g. the coda of the English word [fækt] contains two voiceless stops in a row) or it can be outright defied (e.g. we have no problem with the word [stop] in English, despite [s] being more sonorous than [t]3), and languages differ in how restrictive the sonority hierarchy is. It is also important to note that despite vowels being the most sonorous segments and generally serving as the nucleus, other types of sounds can do so as well. In English it is possible to have glides ([jsp] yup), liquids ([snor.kl] snorkel), and even nasals ([ki?.n] kitten4) can be the nucleus of a syllable.

2.3 Yakima

The bulk of the work done on Yakima has been carried out by Virginia Beavert and Sharon Hargus. They have produced a grammar and dictionary (Beavert and Hargus 2009), as well as several papers (e.g. Hargus and Beavert 2002, 2006), on the language by working with and recording elders in the tribe. Beyond documentation, their investigations have mainly concentrated on phonetics, which while not specifically about syllable structure, necessitates some understanding of it. Hargus and Beavert (2002) give a description of syllable structure in order to examine the phonetics of epenthetic vowels in consonant clusters.

This description begins with a note that there is little previous work on syllable structure in Yakima (and Sahaptin in general) other than the observation that there are frequent consonant clusters. This is a distinctive feature of Sahaptin, as its relative Nez Percé does not allow onset clusters, while both languages allow for coda clusters of up to four consonants. Yakima allows for two consonants of increasing sonority in onsets, with the the ability to affix two additional obstruents, allowing for clusters of up to four consonants.

3 [s] is a common exception to sonority sequencing cross-linguistically, in fact
4 The rest of this transcription is relatively loose, but syllabic [n] does occur in this word
Hargus and Beavert further hypothesize that since \textit{stop + stop + stop} and \textit{fricative + stop + stop} sequences are not attested, it is likely that \textit{stop + stop} sequences are only found word initially. They also observe several place restrictions on phonotactics, summarized in table 1:

<table>
<thead>
<tr>
<th>First Segment</th>
<th>Labial</th>
<th>Coronal</th>
<th>Dorsal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labial</td>
<td>+ [ppaaw] ‘bullsnake’</td>
<td>+ [pčiš] ‘door’</td>
<td>1 [pχ’i-] ‘think’</td>
</tr>
<tr>
<td>Coronal</td>
<td>+ [tpiš] ‘face’</td>
<td>+ [ttáwax-] ‘grow’</td>
<td>+ [tk’i-] ‘look at’</td>
</tr>
<tr>
<td>Dorsal</td>
<td>+ [k’pis] ‘cold’</td>
<td>+ [k’tit] ‘solid, hard’</td>
<td>+ [kki] ‘fringed, drooping’</td>
</tr>
</tbody>
</table>

Table 1: Yakima Onset Clusters

A plus means that a kind of sequence is well attested, while a 1 represents something that is only seen once in the attested forms. They take this data to mean that there is a restriction on \textit{labial + dorsal} sequences specifically and that identical consonants become geminates.

Hargus and Beavert then discuss coda clusters, noting that because of morphological differences, onset clusters are always word-initial, but coda clusters are not necessarily always word final. As stated above, Yakima allows for up to four consonants of descending sonority in final clusters (e.g. [tawq’χς] ‘scarf’). They did not find the same sorts of restrictions on place, or any restrictions at all, aside from sonority (see figure 2).

\[
\begin{align*}
\text{dorsal + coronal} & \quad [q’urχ]\quad \text{‘knee’} \\
\text{coronal + dorsal} & \quad [liχ]\quad \text{‘dirt’} \\
\text{same (gross) place of articulation} & \quad [nuq”k]\quad \text{‘swallow’}
\end{align*}
\]

Figure 6: Yakima Coda Cluster examples

Hargus and Beavert (2006) also have some important insights into syllable structure, specifically about minimality. In this paper they argue that Yakima words are minimally biconsonantal, having either the shape \textit{CCV} or \textit{CVC}. This is somewhat unexpected, since typologically \textit{CV} syllables are quite common, and indeed, they do occur in Yakima, they just must be followed by another syllable (as in [k’úsì] ‘horse’). This pattern cannot be accounted for by requiring disyllables either, since \textit{CCV} words are well attested.

This requirement motivates a greater number of clusters throughout the language, as summarized in this chart:
They did find one word, [ts'a:] ‘size, fit’, that had only one consonant. However, it does have a long vowel (which are often considered to function like an additional consonant), and they point out that few phonological generalizations have truly 100% conformity across the lexicon. This phenomenon, while dealing with whole words rather than syllables, still affects syllable structure, since it motivates a greater number of clusters and is important to account for in an analysis. For instance, a markedness constraint against uniconsonantal words would motivate clusters in GEN. And in fact, they propose such a constraint with the constraint: PrWd $\geq$ CC.

In talking about these patterns of minimality, they also expand further on some of the possible clusters. They note that both obstruent and sonorant initial clusters are possible:

<table>
<thead>
<tr>
<th>cluster pattern</th>
<th>word</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>SN</td>
<td>qni:p</td>
<td>shameful</td>
</tr>
<tr>
<td>SL</td>
<td>ŋaj</td>
<td>pulverized dried salmon</td>
</tr>
<tr>
<td>SJ</td>
<td>twati</td>
<td>medicine person</td>
</tr>
<tr>
<td>FN</td>
<td>tmmm</td>
<td>old woman</td>
</tr>
<tr>
<td>FL</td>
<td>$\chi$'am</td>
<td>large cedar root basket</td>
</tr>
<tr>
<td>FJ</td>
<td>$\chi$jaw</td>
<td>dry</td>
</tr>
</tbody>
</table>

Figure 8: Obstruent cluster patterns in Yakima (S = stop, N = nasal, L = liquid, J = glide, F = fricative)

as well as clusters of same manner obstruents (which in Morelli’s typology constitutes type 6$^5$)

$^5$Hargus and Beavert mention this, but I was unable to find a copy of the paper they referenced, and cannot elaborate on exactly what this means
They find that sonorant initial clusters are more limited, and that they often undergo [i] epenthesis, likely to preserve sonority contours.

And finally, they find examples of identical consonant sequences (as in [tʃuʃaʃa] ‘juneberry’). This is surprising since there is a cross-linguistic pattern, called the Obligatory Contour Principle, that languages tend to disprefer sequences of identical or even very similar phonemes. All together, these descriptions give a solid idea of how syllables can be constructed in Yakima.

Since Yakima and Umatilla are closely related and, in fact, mutually intelligible, this account gives a good place to begin forming hypotheses about Umatilla structures. An initial analysis of the data should look at the size of clusters allowed and any apparent place or sonority restrictions. It would also be beneficial to have a more thorough and focused account of syllable structure, since this description was only a preliminary framework used for conducting a different analysis.

2.4 Umatilla

Umatilla has less published analysis (although more pedagogical materials) than Yakima. The most relevant resources are a sketch grammar and a dictionary (which also contains a grammar). From these there is a solid description of Umatilla’s phonetic inventory, its syntax and inflectional morphology, and a sizable word list (CTUIR and Rude 2014). There are also a few recordings publicly available through the Confederated Tribes of the Umatilla website.

From CTUIR and Rude (2014), we know that Umatilla has a small vowel inventory, but robust consonant one. Voicing is not contrastive, but glottalization and length are. Its morphology is synthetic to mildly polysynthetic, with a robust case system (which draws a human/non-human distinction) that allows for relatively free word order, which is utilized for pragmatic affects. It also inflects
nouns, adjectives, and pronouns for number (singular, dual, and plural), and uses ergative-absolutive alignment. It is noted that Umatilla syllables must have an onset, and those that orthographically appear to be vowel initial actually have a glottal stop at the beginning that, due to convention, is omitted orthographically\(^6\).

Gaps in the literature beyond syllable structure include the stress system, a fuller account of its phonology, and the specific pragmatic meanings of different word orders. Key to understanding the former two is an account of syllable structure, since syllables are the unit to which stress attaches and are often highly relevant to phonological processes.

### 3 Research

During the summer of 2018 I received funding to conduct research with the CTUIR. I had prepared for a year, and in June I began working with the tribe.

#### 3.1 Data Gathered

After getting the appropriate permission, I was able to work with two speakers of Umatilla. From each I was able to get a 52 item word list (a mixture of thirty nine Swadesh items and thirteen additional forms for direct comparison with the Yakima data), four simple sentences, and about four minutes of story telling (using a picture book). While eliciting all but the last of these (the stories were excluded for practicality’s sake), I took notes, specifically marking syllable boundaries and clarifying with my consultants.

One of my consultants also offered additional data in the form of poetry – both original and translated – and translated song lyrics in Umatilla\(^7\). He provided these after I mentioned that I was working on syllable structure, since syllables are highly salient to translating and composing rhythmic work like music and poetry.

#### 3.2 Field Experience

I began attempting to communicate with the CTUIR in mid March, however, it was not until June (when I had made it to the area) that I was able to speak with someone. After a brief phone conversation with the Umatilla Language Program

---

\(^6\)A cursory phonetic investigation upheld this generalization, although a more thorough one may be worthwhile

\(^7\)For full data set, excluding story transcriptions, see Appendix
manager I set up an appointment to visit the Language Program’s offices. There was a small delay due to the death of a tribal elder, but I was able to secure the guidelines for a proposal from the program Director, Modesta Minthorn. After submitting my proposal, I presented it at a meeting of the Education Commission, where it was voted to proceed with the application process. This required some paperwork and a letter of approval signed by the Chairman of the Board of Trustees. After I filled out the forms and secured my letter, I had permission to work with speakers, and after finding a suitable time, did so.

When I arrived I first worked with Thomas Morningowl, a tribal elder and member of the Language Program, who I had met on my previous visit. His mother, Inez Spino-Reves, was Noel Rude’s primary consultant in constructing the Umatilla Dictionary; thus, he was familiar with linguistic field methods, as well as intimately familiar with Umatilla, making him a wonderful consultant. Our recording session took about 15 minutes, and while I was waiting for my next consultant to return from lunch we spoke more generally. It was in this conversation that he told me about his passion for translating and composing songs and poetry in Umatilla. Since such activities rely on syllables for metrical purposes he gave me some copies of his work and allowed me to record him reading some haiku, including some he composed during our conversation. We also discussed some history of the language program and the importance of language in Umatilla belief. It was a great privilege and joy to work with someone so knowledgeable and enthusiastic. My next consultant was Mildred Quaempts, another tribal elder. Our session went similarly well, and she provided some slight variation from Thomas on some items. I ultimately spent less time with her, since she had other things she needed to do, but her help was still very important and appreciated.

This was a complicated process with many steps and it is very fortunate that I anticipated that and started early. As it was, I actually had to delay my return to Swarthmore by a few days to find a time to work with my consultants. Thus, on top of providing me the data necessary for this project, my fieldwork was a powerful object lesson in the importance of planning and initiative.
4 Analysis

4.1 Understanding the Data

4.1.1 Syllable Types

The minimal syllable I found in Umatilla was V (the minimal word being CV), while the maximal syllable in my data was either CCVCC or CVCCC. The following table gives the types, their counts (and percentage of the total in my data), and an example:

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
<th>Percentage</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>25</td>
<td>8%</td>
<td>[i.wa] '3sg-is'</td>
</tr>
<tr>
<td>CV</td>
<td>102</td>
<td>33%</td>
<td>[nu].nu 'nose'</td>
</tr>
<tr>
<td>VC</td>
<td>16</td>
<td>5%</td>
<td>[an] 'sun'</td>
</tr>
<tr>
<td>CVC</td>
<td>87</td>
<td>28%</td>
<td>[tun] 'what'</td>
</tr>
<tr>
<td>CCV</td>
<td>23</td>
<td>8%</td>
<td>[t'kw].ta '1sg-eat'</td>
</tr>
<tr>
<td>CCVC</td>
<td>24</td>
<td>8%</td>
<td>[tpi] 'face'</td>
</tr>
<tr>
<td>CVCC</td>
<td>15</td>
<td>5%</td>
<td>[pip] 'bone'</td>
</tr>
<tr>
<td>CCCV</td>
<td>6</td>
<td>2%</td>
<td>[pxwi] 'think'</td>
</tr>
<tr>
<td>CCVCC</td>
<td>6</td>
<td>2%</td>
<td>[i.wamp] '3.sg.past-is'</td>
</tr>
<tr>
<td>CCVCC</td>
<td>2</td>
<td>1%</td>
<td>[nxokt] 'swallow'</td>
</tr>
<tr>
<td>total</td>
<td>306</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Figure 11: Field Data Syllable Counts

From this it is clear that CV, unsurprisingly, is by far the most common syllable type. While codas can get quite complex, onsets seem to still be generally preferred, with 86% of syllables having at least one onset consonant, compared to only 49% having codas. I did not find any plain V syllables anywhere other than word intially, nor did I find any CCVCC syllables that were in multisyllabic words. These could simply be gaps in my data, perhaps because the bulk of my analysis is on roots and stems without any inflectional morphology, or it could indicate something about how these syllables interact with licensed words.

4.1.2 Patterns and Distributions

Umatilla frequently has sonority plateaus, in both onsets (e.g. [tpi] 'face') and codas (e.g. [nxokt] 'swallow'), indicating that identical manner clusters are allowable. In fact, it seems that most combinations of manner are licensed in onsets as shown in Table 12:
The only manner combinations I could not find attested were *Nasal + Glide* and *Nasal + Nasal* clusters. Notably, in addition to the already mentioned sonority plateaus, this means that there are outright violations of sonority hierarchy allowed, such as *Nasal + Stop* clusters.

### 4.2 Comparison with Yakima

#### 4.2.1 Direct Word Comparison

While the data show that Umatilla generally seems to conform to the same patterns as Yakima, there were, however, some differences. Some were not significant, i.e. completely different or largely unrelated words, but there were a few of the direct comparison words that indicate possible differences in syllable structure.

From this data, which are the words that I was able to gather that were also in the work by Hargus and Beavert, some potential additional limitations that Umatilla imposes seem to emerge. First, Umatilla does not have the *k’p* cluster in “cold” that Yakima does. It seems that this is not due to any phonotactic reason, as we see that cluster in other words such as *k’piit* “ball”. Thus, this difference is likely just lexical and not particularly significant.

Secondly, in the word for ‘dirt’, we see that Umatilla does not allow the same *tx* cluster that Yakima does. However, it is not just that these two segments are not allowed to cluster, because in the earlier word *k’oxl*, we see the reverse order. It could be a restriction on *coronal + dorsal* clusters that Yakima does not have. If so, it must apply specifically in codas, since the word *nkət* “meat” is an example of such a cluster in an onset. However, one of my speakers also produced *t̪ajk* for “dirt”, which includes a coronal and a dorsal together in a coda. This may simply
be a case of variation between speakers, and without more consultants I do not think I can resolve this.

Relatedly, in the word for 'swallow', we see some differences, seeming to stem from disprefering the q"k cluster. My consultants gave me different translations for this word, but both avoided this cluster, resulting in either a kt or xt cluster. What is interesting here is that both elements changed, each moving forward a place in articulation. I actually found q quite rarely in my data, generally in onsets; I only found it once in a coda (*fapawqmu*jpa, “they make them bucks"), where it is mostly likely a singleton coda or else the first segment of an onset cluster. Therefore, I posit that Umatilla disallows q in coda clusters.

Finally, the word for 'meat' is the one case where Umatilla appears more permissive than Yakima, allowing a nk cluster where Yakima epenthesized a vowel. This may be significant, however, it was noted in Hargus and Beavertt (2006), where I am comparing, that there is alternation with this epenthesis.

4.2.2 Structural Comparison

Umatilla and Yakima seem to handle syllable structure very similarly in many ways. First, in figure 12, it was clear that most of the onset combinations allowed in Yakima are also found in Umatilla. Furthermore, in figure 13, most of the direct

\[\begin{array}{|c|c|c|}
\hline
\text{Gloss} & \text{Umatilla} & \text{Yakima} \\
\hline
\text{Face} & \text{tpif} & \text{tpif} \\
\text{cold} & \text{k'sit, k'sat} & \text{k'pis} \\
\text{door} & \text{ptif} & \text{ptif} \\
\text{grow} & \text{twa} & \text{twa} \\
\text{hard} & \text{k'ta, k'as} & \text{k'tit} \\
\text{think} & \text{pxwi, pwit} & \text{p'i-} \\
\text{knee} & \text{k'oxi} & \text{q'uxd} \\
\text{dirt} & \text{q'wx, q'x} & \text{q'uxd} \\
\text{swallow} & \text{nxoxt, nuxt} & \text{nuq"k} \\
\text{dry} & \text{xiul, hiao} & \text{xjaw} \\
\text{be angry} & \text{sxat, saxni} & \text{sxix} \\
\text{meat} & \text{nk'at, nk"at} & \text{nik"it} \\
\text{where} & \text{maan} & \text{minan} \\
\hline
\end{array}\]

Figure 13: Umatilla and Yakima Direct Word Comparison
comparison words were the same or functionally so. I was also only able to find one instance of a labial + dorsal combination in my data, and for the same word even (pxwi ‘to think’). However, there are also some apparent differences.

First, Umatilla clusters seem, in general, to be smaller. The maximal size of both onsets and codas in Yakima was a consonant larger than I found anywhere in my data, specifically four compared to three. Furthermore, the most common syllable type I found was CV, versus Yakima’s CVC. Thus, Yakima seems to both allow and prefer larger syllables than Umatilla.

Related to this, Hargus and Beavert’s claim that the minimal Yakima word is biconsonantal does not hold true for Umatilla. I found numerous words (na -‘do’, na -‘not’, and fi -‘this’) that unambiguously had a single consonant. I suspect that this fundamental difference is at an at least partial explanation for why Umatilla has generally less complex margins.

Finally, Umatilla seems to be more lenient when it comes to violations of the sonority hierarchy. It allows both word final Stop + Stop (e.g. inampt ‘horn’) and Stop + Fricative + Stop (e.g. k’itin ‘tooth’) sequences anywhere, both of which are illegal in Yakima. Further, in the word for ‘meat’ Yakima epenthesizes a vowel between the nasal and the less sonorous stop (Hargus and Beavert 2002), but Umatilla does not. Taken together, these facts indicate that Umatilla adheres less strictly to the sonority hierarchy.

5 Conclusion

Umatilla allows for up to three consonants in either the onset or the coda, with a maximal syllable of either CVCCC or CCVCC. Most syllables have an onset, and the most common syllable shape is CV. Most possible combinations of manner and place are attested, including ones that break the sonority hierarchy or form sonority plateaus. There are some potential positional restrictions, such as disallowing [q] in coda clusters and dispersferring complex onsets of Labial + Dorsal. The minimal word in Umatilla is uniconsonantal, while the minimal syllable is V. It shares most of its syllable structure with Yakima, a closely related and mutually intelligible language, but generally has smaller margins and allows a greater number of sonority violations. Their extreme similarity is not surprising, since they are related dialects and have been in contact for thousands of years.

Having a greater understanding of syllable structure can benefit the effort of language revitalization in two main ways. First, having a greater understanding
of syllable structure allows for further linguistic investigation into prosodic and metrical features of Umatilla, such as stress assignment. More scholarship also generally increases the interest in and ability to do linguistic work on a language, so every contribution to the body of literature makes it easier for that body to grow. Second, understanding syllable structure is potentially practically useful in creating language resources. Speech recognition software, pedagogical materials, and morphological parsers all benefit from such knowledge. Therefore, while the contribution of this paper is small, it has the potential to have lasting benefits to the speech community in preserving their language.

5.1 Further Research

By no means have I resolved all questions relating to syllable structure in Umatilla. For one, this information can be used as a basis for an analysis of stress in Umatilla. There are also things directly related to syllable structure to further investigate. I had some significant divergence in my data between my two speakers, so interviewing more speakers of Umatilla would allow one to have a better picture of the distribution of those alternations. More speaker data would also potentially clarify whether [q] is found in coda clusters anywhere, since it is indicated, but certainly not proven, by my data. Finally, due to the limitations of my time, there was no feasible way to thoroughly analyze the entire Umatilla dictionary. This, combined with the relatively small size of the data set I gathered, means that going through more data may indicate that my analysis was wrong in places (such as the maximal syllable) or that there are more patterns that I simply did not have enough pieces of to see. Regardless, it would certainly be a worthwhile endeavor to do a larger scale investigation of the questions I was pursuing here.
References


6 Appendix

6.1 Elicitation Data

My transcription is often loose, since exact phonetic values were not relevant to my research. Therefore, I am making no particular claims about things like vowel identity.

<table>
<thead>
<tr>
<th>English</th>
<th>Consultant 1</th>
<th>Consultant 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I eat meat</td>
<td>tkʷətəʃ nkut</td>
<td>nk'ut nəʃ kʷətəxə</td>
</tr>
<tr>
<td>This is a new dog</td>
<td>tʃi iwət sumpti k'usi k'usi</td>
<td>tʃi iwə tʃampti k'usi k'usi</td>
</tr>
<tr>
<td>Do you know him</td>
<td>miʃəm a[a kʷa[a</td>
<td>im nə jəkʷəjə shin iwə səkʷə jən im</td>
</tr>
<tr>
<td>We will not go</td>
<td>cha powenata</td>
<td>tʃəo na kʷox winata</td>
</tr>
</tbody>
</table>

Sentence Elicitation Data

<table>
<thead>
<tr>
<th>English</th>
<th>Consultant 1</th>
<th>Consultant 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>blood</td>
<td>taliwə</td>
<td>taliwəl</td>
</tr>
<tr>
<td>bone</td>
<td>pipʃ</td>
<td>pipʃ</td>
</tr>
<tr>
<td>car</td>
<td>weŋtrəa</td>
<td>automobil</td>
</tr>
<tr>
<td>computer</td>
<td>ploswitiwas</td>
<td>plus</td>
</tr>
<tr>
<td>die</td>
<td>k'awə</td>
<td>ik'liawi</td>
</tr>
<tr>
<td>dog</td>
<td>k'usi k'usi</td>
<td>k'usi k'usi</td>
</tr>
<tr>
<td>ear</td>
<td>mʃiʊ</td>
<td>mʃu</td>
</tr>
<tr>
<td>eat</td>
<td>t'kʷətə</td>
<td>t'kʷətə</td>
</tr>
<tr>
<td>egg</td>
<td>tamam</td>
<td>tamam</td>
</tr>
<tr>
<td>eye</td>
<td>aʃəʃ</td>
<td>aʃəʃ</td>
</tr>
<tr>
<td>face</td>
<td>tpiʃ</td>
<td>tpiʃ</td>
</tr>
<tr>
<td>fire</td>
<td>iluxʃ</td>
<td>iluk's</td>
</tr>
<tr>
<td>fish</td>
<td>waikanəʃ</td>
<td>waikanis</td>
</tr>
<tr>
<td>full</td>
<td>ka'akəm</td>
<td>kul, kəkəm</td>
</tr>
<tr>
<td>give</td>
<td>ani’</td>
<td>ani’</td>
</tr>
<tr>
<td>hand</td>
<td>apap</td>
<td>apap</td>
</tr>
<tr>
<td>horn</td>
<td>jukas</td>
<td>i'nampt</td>
</tr>
<tr>
<td>I</td>
<td>in</td>
<td>in</td>
</tr>
<tr>
<td>know</td>
<td>jukwaə</td>
<td>jikʷə</td>
</tr>
<tr>
<td>meat</td>
<td>nkət</td>
<td>nkʷət</td>
</tr>
<tr>
<td>moon</td>
<td>alxaix</td>
<td>alxaix</td>
</tr>
<tr>
<td>English</td>
<td>Oneonta</td>
<td>Wanutja</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>name</td>
<td>onetʃ t</td>
<td>wanuʔtʃ t</td>
</tr>
<tr>
<td>new</td>
<td>tsumaŋpti</td>
<td>tʃumti</td>
</tr>
<tr>
<td>nose</td>
<td>nuʃnu</td>
<td>nuʃnu</td>
</tr>
<tr>
<td>one</td>
<td>naxʃ</td>
<td>naxʃ</td>
</tr>
<tr>
<td>salt</td>
<td>kaˈwanəʃ</td>
<td>sul</td>
</tr>
<tr>
<td>stone</td>
<td>pʃwa</td>
<td>pʃwa</td>
</tr>
<tr>
<td>sun</td>
<td>an</td>
<td>an</td>
</tr>
<tr>
<td>tail</td>
<td>waɬwaʃ</td>
<td>waɬwaʃ</td>
</tr>
<tr>
<td>television</td>
<td>ʃapakˈɪnutpama</td>
<td>ʃəpaˈɪnutawas</td>
</tr>
<tr>
<td>this</td>
<td>tʃi</td>
<td>tʃi</td>
</tr>
<tr>
<td>to think</td>
<td>pxwi</td>
<td>pˈwiʃt</td>
</tr>
<tr>
<td>tongue</td>
<td>məlaʃ</td>
<td>malaʃ</td>
</tr>
<tr>
<td>tooth</td>
<td>kˈʃtin</td>
<td>kˈʃtin</td>
</tr>
<tr>
<td>two</td>
<td>napt</td>
<td>nupt</td>
</tr>
<tr>
<td>water</td>
<td>tʃuʒ</td>
<td>tʃuʃ</td>
</tr>
<tr>
<td>what</td>
<td>tun</td>
<td>tun</td>
</tr>
<tr>
<td>who</td>
<td>ʃin</td>
<td>ʃin</td>
</tr>
<tr>
<td>wind</td>
<td>huli</td>
<td>woli</td>
</tr>
<tr>
<td>year</td>
<td>anuwitʃt</td>
<td>anuwitʃt</td>
</tr>
<tr>
<td>you (singular)</td>
<td>im</td>
<td>im</td>
</tr>
<tr>
<td>you (plural)</td>
<td>amaɫaʃ</td>
<td>—</td>
</tr>
<tr>
<td>cold</td>
<td>kˈʃit</td>
<td>kˈʃet</td>
</tr>
<tr>
<td>door</td>
<td>pʃʃtʃ</td>
<td>pʃʃtʃ</td>
</tr>
<tr>
<td>grow</td>
<td>ʃəpəstʃəx</td>
<td>tətəxt</td>
</tr>
<tr>
<td>hard</td>
<td>kˈtə</td>
<td>kˈwaʃ</td>
</tr>
<tr>
<td>knee</td>
<td>kˈoxʃ̊</td>
<td>koxʃ̊</td>
</tr>
<tr>
<td>dirt</td>
<td>ɬəʃtʃk</td>
<td>ɬəɬ</td>
</tr>
<tr>
<td>swallow</td>
<td>nxokt</td>
<td>nuxt</td>
</tr>
<tr>
<td>dry</td>
<td>xiau</td>
<td>hiao</td>
</tr>
<tr>
<td>be angry</td>
<td>sxətə</td>
<td>saxnį</td>
</tr>
<tr>
<td>where</td>
<td>maan</td>
<td>ma:n</td>
</tr>
</tbody>
</table>

Word Elicitation Data
6.2 Songs

For the songs and poems, since most of them I did not hear and transcribe, but was given written in orthography, I am less confident in the conclusions I draw based on them. However, I treat (and genuinely expect) them to be broadly correct.


\begin{verbatim}
inmipáynam wa čná klawit
Tonight you’re mine completely

panišamá atawit inmiyaw
You give your love so sweetly

łaqayxiša ačašpa imatawit
Tonight the light of love is in your eyes

mišnam wata čná maysxipa
But will you love me tomorrow

miš či iwa kwiyaam, ataw
Is this a lasting treasure

uumiš awtika pakuussa
Or just a moment’s pleasure?

awku payš kwiiknam inmiyay pxwiša
Can I believe the magic of your sighs?

mišnam atawitaas maysxipa
Will you still love me tomorrow?

awčna susutkisimnam nawa
Tonight with words unspoken

iinisim wata imiilaykay
You say that I’m the only one

miš naš wataxna pišilksa
But will my heart be broken

anamuun imayičušmš aan
When the night meets the morning sun?

šukwatatašaaš imiin timna
I'd like to know that your love
\end{verbatim}
Amarillo maytski-yaw, kwnakni San Antone
Amarillo towards the morning, to there from San Antone

Tl'awxna tuun inmisim, ash kuush ashapshasha,
All that which is mine only, I now carry upon my back.

Kawna muun xwaami ich'ushamsh An
For when the sun shines high above

Washataash shapawaqmuypa
I will ride where “the make them buck” (rodeo)

Amarillo maytski-yaw, Amarillo, kwnaxi-taash.
Amarillo towards the morning, Amarillo, there too shall I be.

Pa’paxwiyaash wasat’uus, pashapa k’upnash wixa,
They stole my saddle, they broke my leg,

Anakshanaash tilaki, iwilalakshanaash ilksa.
My wife left me, she left me all alone.

Awashataash kusina maysx, timashataxnaash tsaa
I will ride the horse tomorrow, I wmight be written close,
Amarillo maytski-yaw, Amarillo kwnaxi-taash.
Amarillo towards the morning, Amarillo, there too shall I be.

Amarillo maytski-yaw, kwnakni San Antone.
Amarillo towards the morning, to there from San Antone.

Tl’awxna tuun inmisim, ash kuush ashapshasha,
All that wich is mine only, I now carry upon my back.

kutsk wa ushanikt, inmisim iwa
small amount of what is treasured, it is mine only

chawshwa miyuux, walimnash wa
I am not a chief, only a common man

Amarillo maytski-yaw, Amarillo kwnaxi-taash.
Amarillo towards the morning, Amarillo, there too shall I be.

Amarillo maytski-yaw, Amarillo kwnaxi-taash.
Amarillo towards the morning, Amarillo, there too shall I be.

Song 2- Amarillo

6.3 Poems

Once upon a midnight dreary, while I pondered, weak and weary,
Iwacha pupachu scat, pxwinawashanaash shalawiyaat,
It was midnight, I was tiredly thinking

Over man a quaint and curious volume of forgotten lore
Killushaash miima timash ana tuun iwacha laalakwi shuukwat
I look at old papers, that are of knowledge forgotten

While I nodded, nearly napping, suddenly there came a tapping,
Kpayl nash chuun, chawwiyatnash pnun, kuuk inuun.
for a short time I slept almost lying down, then it sounded

As of some one gently rapping, rapping at my chamber door.
Kuush tuun iwiyawaniya klak, Kawkna pchishpa iinmi.
then what it came “klack”, it knocked at my door
“Tis some visitor,” I muttered, “tapping at my chamber door —
Shin chi iwampsh iinmipayn, kawkna pchish inmii?
who is this that come to be beside me knocking at my door?

Only this, and nothing more.”
Chisim kwaal chawtuul maal.
only this far, nothing more.

Poem 1- The Raven (exerpt)

Wášnaš ín tanán
I am native

Tiičaminšmi míyanaš
A child of the earth

Kʷalísím wátás
I will always be

Poem 2 - Ín ‘me’

Latíša aw xnit
The roots now bloom

Wawáxam iwám̓š cmúy
Springtime comes warm

Ayáyat latít
Beautiful flower

Poem 3 - Xnit ‘roots’

Ituníššamš
It is coming up the river

Áwtni tkʷátat waykánaš
The holy food salmon

Ayáyat čušpa
In the beautiful water
Poem 4 - Waykánaš ‘salmon’

şpawan iganawi
fall has arrived
mijanāša puskwli
the children are at school
pagu qwwaša 'iš
the mothers are happy

Poem 5

miima tilaki
old lady
čimti winš pinmilaykay
new man for her
awapatša
just to help her

Poem 6

inamsa latit
the flowers are going away
špam iwams kpuyl iwata
fall is coming, in a short while it will be winter
anam payu k’sit
very cold

Poem 7