ICA 2013 Montreal  
Montreal, Canada  
2 - 7 June 2013  

Speech Communication  
Session 5aSCb: Production and Perception II: The Speech Segment (Poster Session)

5aSCb5. The Nez Perce vowel system: A phonetic analysis  
Katherine Nelson*

*Corresponding author's address: Linguistics, Rice University, Houston, Texas 77005, katielnelson@rice.edu

The Nez Perce language, a highly endangered American Indigenous language, has been of great interest in phonology over the years due to its unusual vowel system and vowel harmony process. It has five monophthongs and seven diphthongs, all with phonemic length. This system is unusual because rather than /i, e, a, o, u/ as is common, the Nez Perce inventory is /i, æ, a, o, u/. This uncommon inventory leads to two seemingly unrelated dominant, /i, a, o/, and recessive, /i, æ, u/, vowel harmony groups. To date there has been no phonetic analysis of the vowel system. This paper provides an acoustic analysis of the vowels as well as the vowel harmony system. Five native speakers (two male and three female) were used to analyze the vowels and the three female native speakers were used for the vowel harmony study. Results support the current vowel system analysis for Nez Perce. The vowel harmony data lends support to the current advanced tongue root analysis; however it also poses questions for future research.

Published by the Acoustical Society of America through the American Institute of Physics

©2013 Acoustical Society of America [DOI: 10.1121/1.4800241]  
Received 22 Jan 2013; published 2 Jun 2013  
INTRODUCTION

Nez Perce is a highly endangered American Indigenous language of the Pacific Northwest. The language is a member of the Sahaptin branch of the Penutian language family and is most closely related to Sahaptian. Its phonology has been a subject of great interest, due to the language’s unusual vowel harmony system (Aoki 1966; Zimmer 1968; Jacobsen 1968; Rigsby and Silverstein 1969; Zwicky 1970; Silverstein 1979; Hall and Hall 1980; Crook 1999; Bakovic 2000; MacKenzie and Dresher 2003). None of these previous phonological studies, however, have included any acoustic data to support their analyses. The aim of this paper, then, is to provide a phonetic analysis of the vowel system, to support (or question) the previous phonological analyses of Nez Perce, as having a unique vowel and vowel harmony system.

Several previous studies have suggested that Nez Perce has an unusual vowel system: it displays five monophthongs (FIGURE 1) and seven diphthongs (FIGURE 2), all with phonemic length. Previous work has pointed out that this vowel inventory is unusual for a five-vowel system, due to the presence of the low front vowel rather than a mid front vowel (Maddieson 1984; Schwarts, Boë, Vallée and Abry 1997).

Monophthongs

<table>
<thead>
<tr>
<th></th>
<th>Short</th>
<th></th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Front</td>
<td>Central</td>
<td>Back</td>
</tr>
<tr>
<td>High</td>
<td>i</td>
<td>u</td>
<td>i:</td>
</tr>
<tr>
<td>Mid</td>
<td>o</td>
<td>o</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>æ</td>
<td>a</td>
<td>æ:</td>
</tr>
</tbody>
</table>

FIGURE 1. Nez Perce monophthongs (Crook 1999: 21)

Diphthongs

<table>
<thead>
<tr>
<th></th>
<th>Short</th>
<th></th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Front</td>
<td>Central</td>
<td>Back</td>
</tr>
<tr>
<td>iw</td>
<td>aw</td>
<td>uy</td>
<td>i:w</td>
</tr>
<tr>
<td>æw</td>
<td>ay</td>
<td>oy</td>
<td>æ:w</td>
</tr>
</tbody>
</table>

FIGURE 2. List of the Nez Perce diphthongs

Maddieson (1984) finds that of the 317 languages in UPSID (UCLA Phonological Segment Inventory Database), 209 languages display the triangular five-vowel system, /i, ɛ, a, ə, u/, shown in FIGURE 3, and less than five percent of the languages with a five-vowel system use a configuration other than this triangular pattern. The triangular pattern is often seen because languages like to have a maximally dispersed vowel space (Liljencrants and Lindblom 1972). In addition, speakers perceive the phonetic segments they hear as hyperarticulated, or more peripheral in the vowel space (Johnson, Flemming, & Wright 1993). Because languages tend to have maximally dispersed vowels, it is uncommon for there to be a gap, such as the lack of the mid front vowel in Nez Perce. Furthermore, only 43 of the 317 languages, or 13.6%, have a gap in their inventory (Maddieson 1984: 136).
If a language does have a gap in its vowel inventory, this missing vowel is least likely to be the high front vowel, /i/, or the low central vowel, /a/, suggesting the high back vowel is more likely to be missing (Maddieson 1984: 142). Maddieson also finds that the general hierarchy for a five vowel system is \{i, a\} \{e, o\} \{u\} (142). Nez Perce, however is lacking /e/ and does not follow this hierarchy, though Rigsby and Silverstein (1969) suggest that the low front vowel /æ/ developed from a mid front vowel, /e/.

The gap seen in the Nez Perce inventory is caused by a vowel of unexpected height (Maddieson 1984: 147-148). Maddieson finds that in the 317 languages there are only three languages with this gap: Nez Perce, Taishan and Ket. All of these languages use /æ/ in place of /e/ or /ɛ/ (Maddieson 1984: 148).

In addition to the vowel inventory outlined above, Nez Perce has a vowel harmony system that affects all of the vowels. The vowel harmony groups are outlined in **FIGURE 4**, showing the dominant and the recessive vowels sets.

<table>
<thead>
<tr>
<th>Dominant</th>
<th></th>
<th>Recessive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front</td>
<td>Central</td>
<td>Back</td>
</tr>
<tr>
<td>High</td>
<td>i</td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>a</td>
<td>o</td>
</tr>
<tr>
<td>Low</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 4.** The dominant and recessive vowel inventory (Crook 1999: 245).

Nez Perce vowel harmony has predominately been described in terms of ATR or RTR (advanced or retracted tongue root) (Jacobsen 1968; Rigsby and Silverstein 1969; Zwicky 1971; Hall and Hall 1980; Bakovic 2000; MacKenzie and Dresher 2003), though the system does not appear to fit an ATR analysis; many of the features seen in ATR vowel systems are not seen in the Nez Perce vowels. The dominant vowels are usually +ATR in an ATR vowel harmony language (Hall and Hall 1980: 220-221; Casali 2003: 307), however if Nez Perce does indeed display ATR vowel harmony, the dominant set is the –ATR vowels while the recessive set is the +ATR vowels (Hall and Hall 1980: 220-221). Furthermore, ATR vowel harmony languages with a neutral vowel often show this neutral vowel in the low vowel (Ohala 1994: 493). Nez Perce, however, has a vowel harmony pair /a/ and /æ/ in the low position and a single high front vowel that is present in both the dominant and recessive sets. This high front vowel looks like a neutral vowel but does not behave as a neutral vowel would. This vowel can condition vowel harmony and is a member of both the dominant and recessive vowel sets. Previous studies either do not mention an acoustic difference between the dominant and recessive [i]’s or say that they are not acoustically distinct (Crook 1999: 245). Though the high front vowel is not a neutral vowel it is unusual to have a single vowel in the high front position while there is a vowel harmony pair in the low position.

The final vowel harmony pair is the mid back and high back vowels. Unlike other vowel harmony pairs these vowels are not of the same height. As will be shown, however, the /o/ is produced rather high making the vowels closer in height.
To date Nez Perce vowels have not been instrumentally studied and plotted to determine the acoustic vowel space. This has led to, for instance, scholars using different IPA symbols for the same vowels, such as transcribing the mid back vowel as both [o] and [ɔ] (Bakovic 2000; MacKenzie and Dresher 2003). This acoustic study will provide further data for understanding the unique Nez Perce vowel system and vowel harmony. Additionally, it will also provide information for the theory of maximal dispersion (Liljencrants and Lindblom 1972) and the hyperspace effect (Johnson, Flemming, & Wright 1993). While the Nez Perce vowels do not appear to follow the theory of maximal dispersion, the following analysis of Nez Perce suggests that they are indeed maximally dispersed but not in the expected way. The theory of maximal dispersion also suggests an alternative analysis to the current ATR analysis for vowel harmony.

**METHOD**

Data for this paper were collected during two trips to Lapwai, Idaho during the summers of 2010 and 2011. Word lists were used to elicit the data. Words for the lists were collected from various language-learning materials, the Nez Perce Dictionary (Aoki 1994), previous publications (Crook 1999), and native speakers. Words chosen from these texts were examined by native speakers and archaic words were eliminated. In addition any ungrammatical constructions were corrected or eliminated. Words deemed suitable were compiled into a final word list. The words used in the vowel harmony study had the structure of stem-stem or stem-suffix. To facilitate comparisons, noun-suffix paradigms were predominantly used. The stem-stem and stem-affix combinations used were of the following structures: dominant-dominant, dominant-recessive, recessive-dominant, and recessive-recessive.

Five native speakers were recorded for this study. Additionally, there are two sets of data in this study, one set for the vowel system in general and another set specifically for vowel harmony. The exact number of tokens varies person to person, as some speakers were able to give more time than others. Tokens were discarded due to mispronunciations, background noise, devoiced vowels (for the target vowels), or if upon analysis one of the aspects of study was not calculable; for example, formants, bandwidth or if there were outliers. These token counts include words that were repeated three times. For the formant analysis (for both the vowel harmony and vowel data), these repetitions were averaged together for each speaker. Some words included multiple tokens used in the analysis. For the bandwidth and harmonics analysis (for the vowel harmony data) the tokens were averaged by token type rather than by the set of three repetitions.

Speakers were recorded in quiet rooms using a Zoom H4n recorder and an AKG C555L head-mounted microphone with an AKG MPA VL adapter. Speakers were prompted with the English translation and responded in Nez Perce. They were then asked, in Nez Perce, to repeat the word in a Nez Perce carrier phrase or in isolation. Speakers said the word two or three times depending on time constraints.

All statistical analyses were independent sample t-tests. The t-tests were conducted on the vowel harmony pairs, dominant and recessive /i/, /a/ and /æ/, as well as /o/ and /u/, to determine if there are any statistically significant differences between any of the formants, bandwidths, or harmonics.

**RESULTS**

The results are separated into three sections. The monophthongs are presented first with a discussion of the theory of maximal dispersion, followed by a short discussion about the diphthongs, and finally the vowel harmony section. In the monophthong and diphthong sections, Nez Perce is discussed as a five-vowel system including the dominant and recessive high front vowels as one unit. In the section on the vowel harmony system the dominant [i] and recessive [ɪ] are discussed independently.

---

1 All previous work has been done by ear.
2 Though carrier phrases are intended to elicit more natural speech I was told after using a carrier phrase, “Héenek’e _____ híce”, that carrier phrases were in fact not very natural for these speakers, as the words were not being used in a conversation or story.
3 Not all speakers had enough time to repeat the words three times, so to ensure each word on the word list was collected some speakers only repeated the words twice.
Monophthongs

The results of this study support previous phonological analyses of the vowel system. **FIGURE 5** shows the vowel space for the two men in the study and **FIGURE 6** shows the vowel space for the three women.\(^4\) **FIGURE 5** and **FIGURE 6** show that long vowels are more peripheral to the short vowels for all speakers.\(^5\) The [o] is very high and many tokens are of the same height as [u]. The long and short high front vowel and the long and short low back vowel display the largest difference between the long and short monophthongs. The long [a:] is backed and slightly lowered in comparison to [a] while [i:] is raised and fronted in comparison to [i]. Additionally, [i] is produced similarly to [i]. Generally, the vowel spaces for [u:] and [u] and for [o:] and [o] overlap with the short vowels slightly more central in the vowel space.

**FIGURE 5.** Vowel space for men. The data are not normalized.

**FIGURE 6.** Vowel space for women. The data are not normalized.

---

\(^4\) The data are not normalized, due to the low number of speakers. Data are discussed on a speaker by speaker basis.

\(^5\) Long and short vowels displayed a minimum average difference of 45 milliseconds (MS2, male speaker 2) and maximum average difference of 117 milliseconds (FS3, female speaker 3).
The canonical idea of the theory of maximal dispersion for a five-vowel system is a vowel system composed of /i, e, a, o, u/ forming the triangular vowel space seen in FIGURE 7 (a). However, if we consider the Nez Perce vowel space as seen in FIGURE 7 (b), it shows a similar triangular vowel space, but turned about 90 degrees to the left.

![FIGURE 7](a) Vowel space outlined for the canonical five-vowel system (repeated for the reader’s convenience). (b) Vowel space outline for Nez Perce vowels.

An initial analysis and study of the Nez Perce vowels suggested that they violated the theory of maximal dispersion, however, upon further analysis the vowels are spread in a maximally contrastive pattern. The vowels are organized in the expected triangular shape, but the triangle is turned.

**Diphthongs**

As previously mentioned Nez Perce has seven diphthongs with phonemic length. These diphthongs can be seen in FIGURE 8, which displays the diphthongs for one female speaker (FS1, female speaker 1).

![FIGURE 8](Diphthongs for one female speaker)

The diphthongs with a [w] off glide show trajectories toward the back of the mouth while diphthongs with a [y] off glide show trajectories toward the front of the mouth. These off glides affect the production of the nucleus. For

---

6 The long and short diphthongs did not show as strong a durational difference as the monophthongs. The shortest average difference between the long and short diphthongs is seven milliseconds (MS2) and the longest average difference between the long and short diphthongs is 41 milliseconds (FS3).
example, the [æ] in [æy] is produced more like [ɛ] and the [a] in [ay] is raised and fronted. Furthermore, Crook (1999: 26) noted this variation and wrote that this effect on the nucleus by the off glide is not as strong with long diphthongs. Generally diphthongs with a long nucleus are more peripheral to those with short nuclei. Some unstressed diphthongs were used in this section of the study because no stressed variants were found. The results should be representative of the Nez Perce diphthongs, however stressed exemplars of all diphthongs should be found (if possible) and added to this study.

Vowel Harmony

Various acoustic aspects of the vowels were studied to provide evidence for classifying Nez Perce as an ATR language or to provide evidence for an alternative analysis. Only monophthonal vowels were used in this section of the study. Both stressed and unstressed vowels were used. As will be discussed, stressed vowels (SV) sometimes provided stronger data while both the stressed and unstressed vowels (SUV) sometimes provided stronger data, therefore both sets of data will be discussed. To study the vowel harmony the vowels were plotted and t-tests were run on various aspects of the vowel harmony pairs ([a]-[æ], [o]-[u], and the dominant and recessive high front vowels). The first three formants, harmonic difference (A1-A2), first bandwidth and second bandwidth were studied.\(^7\)

All speakers showed a statistically significant difference in the second formant between [a] and [æ] for both the SUV and SV groups (\(p>0.001\)). Furthermore, all speakers show a statistically significant difference in both the first and second formants for the [o]-[u] pair. The difference in the second formant is significant at \(p>0.001\) for all speakers for both the SUV and SV groups showing that the [o] is statistically significantly back in comparison to [u]. The first formant is also significant for both the SUV and SV groups, however the significance is stronger for the SUV group for FS1 (\(p=0.009\)) and FS2 (\(p=0.003\)). The groups are equally significant for FS3 (\(p>0.001\)). These results show that [o] is statistically significantly lower than [u]. None of the results for the dominant and recessive [i]’s are significant, however the F1 is equal to or higher for the recessive [i] than the dominant [i] and the F2 is higher for the recessive [i] than the dominant [i].

\[\begin{array}{c|c|c}
\text{Dominant} & \text{Recessive} \\
\hline
\text{FS1} & \text{FS1} \\
\text{FS2} & \text{FS2} \\
\text{FS3} & \text{FS3} \\
\end{array}\]

\(F_1\) and \(F_2\) values for vowel pairs

**FIGURE 9.** Vowel plot showing the dominant and recessive [i]’s for all three speakers.

**FIGURE 9** shows the dominant and recessive [i]’s for the three female speakers; the dominant [i]’s are lower and backed in comparison to the recessive [i]’s. This pattern is seen in all vowel harmony pairs. The dominant vowels in the pairs, [i\(_\text{dom}\), a, o], are all backed and lowered in comparison to the recessive vowels in the pairs [i\(_\text{rec}\), æ, u]. (See the previous paragraph for the data supporting [a, æ, o, u].)

The analysis of bandwidths and amplitude differences did not provide conclusive evidence supporting an ATR analysis for the vowel harmony system. There is no consistent pattern in the data, though this variation may be due

---

\(^7\) These features have been found to distinguish ATR vowels in other languages (Lindau 1979; Hess 1992; Fulop, Kari and Ladefoged 1998; Anderson 2003; Guion, Post and Payne 2004)
to the low number of speakers. The second bandwidth is larger for the recessive vowels (+ATR) than for the dominant vowels (–ATR) in nearly all instances except for one vowel pair, [a] and [æ], for one speaker, FS2. The larger bandwidth for the recessive vowels is opposite of what would be expected, which suggests an alternate analysis to ATR harmony. Additionally, there is little significant data in the first bandwidth, though the recessive vowels show a narrower bandwidth than the dominant vowels in all cases. Most importantly FS2 showed significance in the dominant and recessive [i]’s for the first bandwidth. The most reliable distinction between the dominant and recessive vowels is the harmonic difference. The dominant vowels show a smaller harmonic difference than the recessive vowels. This pattern is seen in all the vowels except for the dominant and recessive [i]’s.

Considering the results for harmonic difference and bandwidths, which do not provide conclusive evidence supporting or rejecting an ATR analysis, the theory of maximal dispersion (Liljencrants and Lindblom 1972) suggests an alternate analysis. Both the dominant set [i\text{dom}, a, o] and recessive set [i\text{rec}, æ, u] of vowels are maximally dispersed. FIGURE 10 shows both the maximal dispersion of the vowels and that the dominant vowels are retracted in comparison to the recessive vowels.

\[ F_2 \text{ (Hz)} \]

\[ F_2 \text{ (Hz)} \]

**FIGURE 10.** Comparison of dominant to recessive vowels.

In addition to the maximal dispersion of the vowel space, the other defining feature between the dominant and the recessive vowels is that the dominant vowels appear to be more tightly clustered than the recessive ones (FIGURE 9).

**DISCUSSION AND CONCLUSION**

The phonetic data presented in this paper supports the earlier phonological descriptions of Nez Perce vowels. The language had previously been described as a language with a gap in its inventory (Maddieson 1984) and the vowel space does not fit the canonical idea of maximal dispersion. This study proposed that the language is indeed maximally dispersed and does not have a gap in its inventory. Though the vowels do not appear to be maximally dispersed because there is no mid front vowel but instead a low front vowel, there is maximal dispersion in the vowel system but on a different axis. The triangular vowel space expected in a five-vowel system is present in Nez Perce but the triangle is turned allowing the high front vowel to be the point of the triangle rather than the low vowel. Results for the formants, bandwidths, and harmonic differences did not provide conclusive evidence supporting or rejecting an ATR analysis for Nez Perce vowel harmony, however, an analysis based on maximal dispersion and hyperspace suggested an alternate analysis. Additionally, this study showed that the dominant vowel group is retracted in comparison to the recessive group FIGURE 10, providing further evidence for an alternate analysis to ATR. The dominant vowels are also more tightly clustered than the recessive vowels indicating a dependence on the hyperspace effect. Vowel harmony in Nez Perce has no doubt changed over the years and may have originally been based on ATR but has possibly come to depend on maximal dispersion and the hyperspace effect to differentiate the dominant and recessive vowel groups.
REFERENCES