

Phonetics of Southern Ute vowels

Viktor Kharlamov

Department of Languages, Linguistics & Comparative Literature,
Florida Atlantic University, USA
vkharlamov@fau.edu

Stacey Oberly

Southern Ute Indian Montessori Academy, USA
stacey.oberly@gmail.com

This paper presents an acoustic study of the vowel system of Southern Ute, a Southern Numic Uto-Aztecan language spoken in southwestern Colorado. Previous auditory accounts proposed an inventory of five vowel phonemes that participate in three allophonic processes and contrast in length and stress. We investigate how the vowels are realized at the phonetic level by analyzing F1, F2, duration, spectral emphasis and f0 in over 6000 vowel tokens produced by eight fluent speakers (four female and four male). Our findings provide new phonetic detail for the earlier non-instrumental descriptions of the language, including both expected and previously unreported effects. We confirm the existence of five distinct phonemic categories but show that their prototypical phonetic realizations and positional allophones do not always match the earlier descriptions. We also describe how phonemic length and stress are marked in Southern Ute.

1 Introduction

In this paper, we investigate the phonetic realization of vowels in Southern Ute, a Native American language that belongs to the Southern Numic branch of the Northern Uto-Aztecan family. Southern Ute is most closely related to Southern Paiute, Chemehuevi, and Kawaiisu (Simons & Fennig 2017). It is a severely endangered language that is only spoken fluently by approximately 40 out of around 1400 tribe members (Oberly et al. 2015). The fluent speakers are all tribal elders. Many of them are in their eighties or older. They reside on the Southern Ute reservation in the town of Ignacio and neighboring areas in southwestern Colorado. Most of the speakers are not literate in Southern Ute, which has an official orthography (Givón 2011) as well as an alternative writing system (Charney 1996). In our study, only one of the eight speakers had working knowledge of either orthography.

Similar to other Native American languages, published and unpublished work on the grammar of Southern Ute is exceedingly rare. It consists mainly of three dictionaries (Goss 1961, Charney 1996, Givón 2016), a grammar book (Givón 2011), a doctoral dissertation (Oberly 2008), and one journal article (Oberly & Kharlamov 2015).¹ Two of the dictionaries (Charney 1996, Givón 2016) and the grammar (Givón 2011) provide limited summaries of the sound inventory that are based primarily on Givón's impressionistic transcriptions made during conversations with tribal elders in the 1970s and 1980s. Charney (1996) and Givón (2011, 2016) also focus heavily on orthography as each author advocates a specific writing system. The dissertation (Oberly 2008) presents a pilot phonetic study of the Southern Ute

¹ Givón (2011) and Givón (2016) are revised versions of Givón (1980) and Givón (1979), respectively.

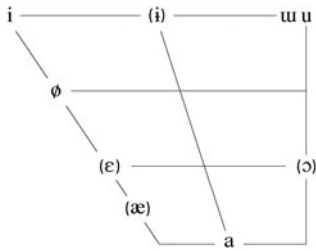


Figure 1 Southern Ute vowels (based on Givón 2011, 2016; allophones are in parentheses).

Table 1 Southern Ute consonants (based on Givón 2016; allophones are in parentheses).

	Bilabial	Labio-dental	Dental	Palatal	Velar	Uvular	Glottal
Plosive	p		t		k	(q)	ʔ
Affricate				tʃ		(qχ)	
Nasal	m		n				
Flap			r				
Fricative		v	s		(x) ɣ	(χ) (ʁ)	
Approximant	w			j			

sound system that is based on a subset of the dataset used in the current study (selected tokens from four of the eight speakers). The only previously published instrumental phonetic study (Oberly & Kharlamov 2015) addresses a specific phenomenon of vowel devoicing and does not examine other patterns or contrasts.

1.1 Southern Ute sounds

The most detailed non-instrumental accounts of the sound inventory of Southern Ute can be found in Givón (2011, 2016). As shown in Figure 1, Givón identifies up to nine vowel sounds. This includes five phonemes, /i u ø a/, and three or four allophones, [ɛ ɔ æ i].² The consonantal inventory (which is relevant to the discussion of vowel allophony) is provided in Table 1 (based on Givón 2016). Other sources, such as the preface to Charney’s dictionary (Charney 1996), do not distinguish between phonemes and allophones and do not provide phonetic descriptions for the vowels. Such sources only list similar-sounding words from North American varieties of English or other languages, which makes it difficult to determine the exact backness and height for the vowels they identify.

For the phonemes, Givón (2016) states that the /i/ is a close front vowel that is ‘roughly like the English vowel in “see”, “please” or “Louise”’ (p. 4). The /u/ is a close back unrounded vowel with no English counterpart. The /u/ is a close back rounded vowel that is ‘roughly like the English vowel in “stupid”, “shoot” or “rouge”’ (p. 5). The /ø/ is a mid front rounded vowel that ‘can be seen in French, as in *feu* (“fire”), or in German, as in *moeglich [sic]* (“possible”)’ (p. 6). The /a/ is an open central vowel that is ‘roughly like the English vowel in “father”, “car”, “John” or “top”’ (p. 5). The similar-sounding English words that are mentioned in Charney (1996) include *police* for the /i/, *number* for the /u/, *boot* for the /u/, *her* for the /ø/, and *father* for the /a/. Examples of Southern Ute words containing the five

² Givón (2011, 2016) uses non-standard transcription symbols that were converted to IPA in all tables and examples. The [i̠] allophone appears in Givón (2011) only.

vowel phonemes can be seen in (1) below (based on Givón 2016). The orthographic form (in italics) is written in Givón's alphabet in which the glottal stop is represented with a vertical line (|), irregular stress is marked with an acute accent (´) placed above the stressed vowel, and voiceless vowels are underlined.

- (1) a. *ʼuni ʼni* /ʔuˈniʔni/ ‘doing/living’
 b. *ʼtʰtʰ* /ʔtʰ:/ ‘yes’
 c. *kamuchi* /kaˈmutʃi/ ‘hare’
 d. *ʼθøvi* /ʔθø:vi/ ‘bone’
 e. *mamachi* /maˈmatʃi/ ‘woman’

For the allophones, Givón (2016) describes the mid back [ɔ] as an allophone of /ø/ that is used before and after velar and uvular consonants and ‘is pronounced roughly as in the English words “door”, “four” or “more”’ (p. 6). Charney (1996: vii) states that the allophony is triggered by adjacency to velars and compares the allophone to the English vowel in *note*. This allophonic pattern is reflected in both orthographies of the language. The official system in Givón (2011) uses the graphemes <ø> ~ <o>. The alternative system in Charney (1996) uses <ö> and <o>.

Givón (2016) also identifies two allophones for the /a/, the mid front unrounded [ɛ] and the low front [æ]. They appear ‘in proximity to the front vowels /i/ or /ø/ or the glide /j/, most commonly preceding but sometimes following /a/’ (p. 5). The [ɛ] is supposed to be favored by younger speakers who ‘pronounce it as in the English words “men” or “get”’ (p. 5). The [æ] is presumably used by older speakers and it ‘resembles more the vowel in English words “cat”, “nap” or “man”’ (p. 5). Charney (1996) only identifies one allophone that occurs next to [i] and [j] and compares its pronunciation to that of the English vowel [ɛ] in *bet*. The /a/ ~ [ɛ/æ] difference is not reflected in the tribe's official orthography (Givón 2011) that only uses the grapheme <a>. However, it is shown in the system in Charney (1996) that uses both <a> and <e>.

The earlier analysis in Givón (2011) also includes the allophone [i]. It is described as a high central vowel that is a positional variant of /u/. The allophone is used in unstressed syllables after non-back (labial, dental or palatal) consonants, which happens ‘typically, but not consistently’ (p. 17). The [i] is no longer listed in Givón (2016) and it is not mentioned in Charney (1996), which suggests that it may not in fact be part of the vowel system. A general summary of the allophonic patterns and examples provided in Givón (2011, 2016) can be found in Table 2.

Southern Ute vowels can also be phonemically short or long. According to Givón (2016), the length contrast affects all vowels and it is independent of stress. Length is reflected in both orthographies by using either a single or a double vowel grapheme (e.g. <a> /a/ vs. <aa> /a:/). There is no phonetic description of what constitutes a short or a long vowel in Givón (2016) but Charney (1996: iv) states that long vowels are ‘twice as long’. An example of a length-based minimal pair is provided in (2) (based on Givón 2016).

Table 2 Southern Ute vowel allophony (based on Givón 2011, 2016).

Phoneme	Allophone	Environment and notes	Examples
/ø/	[ɔ]	Before and after velar/uvular consonants	<i>qhoqʰ</i> /køk/ → [qɰɔqɰ] ‘bull-snake’
/a/	[ɛ], [æ]	In proximity to /j i ø/; younger speakers: [ɛ], older speakers: [æ]	<i>ʼinikjætʰ</i> /ʔiˈnikjætʰ/ → [ʔiˈnikjætʰ]. [ʔiˈnikjætʰ] ‘made’
/u/	[i]	After labial/dental/palatal consonants in unstressed syllables; described as an optional process only in Givón (2011)	<i>tʰijʰ</i> /tuˈijʰ/ → [tiˈiji] ‘deer (OBJ)’

- (2) a. *máykya* [ˈmajkja] ‘said’
 b. *maaykya* [ˈmaːjkja] ‘found’

Furthermore, the vowels of Southern Ute can be stressed or unstressed. According to both Charney (1996) and Givón (2016), two stress patterns are possible: (i) stress on the second syllable (most common) and (ii) stress on the first syllable (rare and marked in both orthographies with the acute accent placed above the vowel). Long vowels are not discussed explicitly but the examples in the dictionaries show that initial syllables with long vowels are always stressed. Givón (2016) or Charney (1996) do not provide any detail on how stress is realized phonetically or how it interacts with pitch accent. Examples of stress placement can be found in (3), including a stress-based minimal pair (3a, b) and a word with a long vowel in the stressed first syllable (3c) (based on Givón 2016).

- (3) a. *paɣʷ* [paˈɣʷ] ‘fish’
 b. *páɣʷ* [ˈpaɣʷ] ‘trout’
 c. *ʼaavʷm* [ˈʼa:vʷm] ‘my upper arm’

Finally, Southern Ute vowels can be devoiced. The process of vowel devoicing is found in several Native American languages of the Rockies and the Plains, including other Numic languages (e.g. Comanche; see Charney 1993). In Givón (2016: 8), devoiced vowels are described as ‘silenced or whispered’. Charney (1996: x) states that they are ‘expressed with a puff of air’. The acoustic study in Oberly & Kharlamov (2015) reveals that their phonetic realization is more complex. Devoiced vowels can be fully voiceless, partially devoiced, or fully reduced/deleted but with a measurable effect on the duration, intensity and voicing of the preceding consonant. An example of a word containing a devoiced vowel in the second syllable is given in (4) (based on Givón 2016):

- (4) *pácha* [ˈpatʃ̥a] ‘shoe’

1.2 Goals and structure

The goal of the current study is to provide a phonetic description of Southern Ute vowels that can be used as a baseline for future research as well as development of pedagogical materials. This is part of a larger endeavor to document language use by the remaining fluent speakers in order to assist with language revitalization efforts (see Oberly et al. 2015). We investigate (i) the general vowel space of Southern Ute, (ii) allophonic alternations affecting vowels, (iii) differences between phonologically short versus long vowels, and (iv) differences associated with stress, including the question of whether stressed syllables carry pitch accent. We address these questions by examining the acoustic measures of F1 and F2 that signal vowel quality and can therefore be used to investigate the general vowel inventory, vowel allophony, and stress-related vowel reduction (see Lindblom 1963, Gay 1978). We also use the measures of duration, spectral emphasis and f0 to examine vowel quantity, stress, and pitch accent (see Fry 1955, Lieberman 1960, Lehiste 1970, Beckman 1986, Traunmüller & Eriksson 2000). Since the phonetic aspects of vowel devoicing in Southern Ute have previously been investigated in detail (Oberly & Kharlamov 2015), in this study we focus specifically on the non-devoiced vowels.

In the rest of the paper, we provide an overview of the methodology and present the results for F1 and F2, vowel duration, spectral emphasis, and f0. We follow up with a general discussion and conclusions on how our findings relate to the previous accounts of Southern Ute, descriptions of related Uto-Aztecan languages, and the literature on vowels in general.

Table 3 Speaker information.

Female	Age	L2/L3	Literate in Southern Ute	Male	Age	L2/L3	Literate in Southern Ute
F1	72	E, S	No	M1	68	E	No
F2	72	E	No	M2	78	E	No
F3	72	E	Yes	M3	79	E, S	No
F4	84	E	No	M4	90	E, S	No

L2/L3 = second/third language, E = English, S = Spanish.

2 Method

2.1 Participants

Our findings are based on recordings from eight fluent Southern Ute speakers, including four females and four males. At the time of recording, their ages ranged from 68 years to 90 years (mean age 76.9 years). All eight speakers acquired Ute as their first language. They all learned English later in life, and three speakers also indicated learning Spanish as an L3. Only one speaker was literate in Southern Ute. A summary of speaker characteristics is provided in Table 3.

2.2 Materials and procedures

Speech tokens came from a translation task, which made it possible to elicit data from the speakers who were not literate in the language. Speakers were presented with 202 English words or short phrases from the list in Oberly (2008) that contains various sound sequences possible in Southern Ute.³ The words were presented orally one word at a time, and speakers were asked to translate them into Southern Ute and produce them both in isolation and in a carrier sentence (*Maas taani . . . máyku* /^hma:s 'ta:ni . . . 'majku/ 'He always says . . .'). Sample stimulus items are provided in (5).

- (5) a. ^hakapi /^hʔa'kapi/ 'buffalo berry (OBJ)
 b. ^hchə'wa /^hʧə'ʔwa/ 'pick up/gather'
 c. ^hisivɰ /^hʔisivɰ/ 'squawberry (SUBJ)
 d. ^hkámaaywatɰ /^hkama:jwatɰ/ 'blind person'
 e. ^haapachi /^hʔa:patʃi/ 'boy (OBJ)
 f. ^hsəθ'mi /^hsə:ʔmi/ 'sip'

The list includes items with the regular stress placement on the second vowel when the first vowel is short (5a, b), irregular stress on the initial short vowel (5c, d), and regular stress on the initial long vowel (5e, f). The full list of stimulus items can be found in Oberly (2008).

Recording took place indoors on the Southern Ute reservation in Colorado. Speakers' responses were captured in .wav format at a rate of 16 bits and with a sampling frequency of 44,100 Hz using a Marantz PMD671 portable solid state digital recorder and a Countryman Isomax E60P5T omnidirectional head-mounted microphone with a frequency response of 30 Hz to 20 kHz. Recording settings and microphone distance were kept constant across the sessions.

³ Southern Ute is an agglutinative language, so speakers sometimes produced wordforms with additional suffixes added onto the stem as they were more familiar with such forms than dictionary-type definitions. They also occasionally produced a related word instead of the expected form or repeated the words more frequently than instructed. All such items were transcribed and included in the analyses, so that the dataset used for statistical modelling would contain as many vowel tokens as possible.

Table 4 Vowel tokens: summary.

Vowel	Number of tokens				
	Total	Short	Long	Stressed	Unstressed
/i/	1504	1354	150	471	1033
/u/	627	575	52	171	456
/ø/	637	506	131	236	401
/a/	538	503	35	233	305
	2889	2397	492	1096	1793
Totals	6195	5335	860	2207	3988

2.3 Data analysis

As shown in Table 4, a total of 6195 vowel tokens were included in the study. Segmentation and measurements were done with Praat (Boersma & Weenink 2017). Vowel boundaries were determined manually on the basis of formant structure (i.e. beginning and end of a clear F2). We used only those tokens where segmental boundaries could be identified unambiguously and only analyzed the vowels in non-devoicing environments.

A Praat script was written to obtain five acoustic measures for each vowel: F1, F2, duration, spectral emphasis, and f0. F1 and F2 were measured in Hz at the vowel mid-point and transformed to *z*-scores, which scales the original values to compensate for between-speaker differences in vocal tract morphology (for plotting and statistical testing purposes; see Lobanov 1971). Vowel duration was measured in milliseconds, followed by log-transformation (for statistical tests; see Chladkova, Escudero & Boersma 2011). For spectral emphasis, which is a measure of vocal effort that correlates with stress (see Eriksson et al. 2016), the study used the method described in Traunmuller & Eriksson (2000). It is based on calculating the difference between (i) the overall intensity of the vowel and (ii) the intensity in a low-pass-filtered signal (with a cutoff frequency of 1.5 f0), which produces a normalized measure of energy in the higher frequency bands. Prior to measuring intensity, the individual word tokens were normalized on the basis of the speakers' average intensity during the entire recording session. For f0, the vowel's mean fundamental frequency was measured over the middle 50% of the vowel. The measurement was done in Hz (for illustrative reasons) as well as ERB (for statistical testing; see Moore & Glasberg 1983, Glasberg & Moore 1990).

Plotting and statistical analyses were done in R (R Core Team 2017). Vowel plots show the F1 × F2 vowel space (using transformed values). The plots were generated with the 'phonR' package (McCloy 2016). Boxplots show the median, the interquartile range and the maximum and minimum values (excluding outliers). They were created using the standard R 'boxplot' function. Inferential statistics were done using the 'afex' package (Singmann et al. 2017) that fits data in mixed models with *lmer* and calculates *p*-values using parametric bootstrapping. Dependent variables were the (transformed) measures of F1, F2, duration, spectral emphasis, and f0. Fixed effects differed depending on the measure and the research question and included such factors as vowel identity (/i u ø a/), lexical stress (stressed, unstressed), phonemic length (short, long), preceding and following environments (allophony trigger present vs. absent), word-finality (non-word final vs. word-final), speaker gender (female, male), and two-way interactions of the fixed effects. Speakers and words were always entered as random effects (including by-speaker slopes, when applicable; see Barr et al. 2013). We also added speech type (citation form vs. carrier sentence) as a random effect to account for the differences between the two contexts.⁴ Model selection process always started with

⁴ The carrier sentence tokens were not truly representative of connected speech as speakers tended to produce the target words similarly with and without the carrier frame (with falling intonation, occasionally

a model containing all relevant fixed effects and the full random effects structure. If the full model did not converge, this was followed by reducing the random effect structure (starting with the item structure). The maximal models that converged are described in the Results section. Pairwise comparisons were done using the least-squares means method with Bonferroni adjustment for multiple comparisons using the ‘emmeans’ package (Lenth 2018).

3 Results

3.1 F1 and F2

Results for F1 and F2 are illustrated in Figure 2, which presents z-score-transformed group means (2A) and individual speaker means (2B) for the phonemes /i u ũ ø a/ as identified in Givón (2016). Individual variability is shown with ellipses that correspond to \pm one standard deviation.

As reflected in Figure 2, mean F1 was low for the three close vowels (/i/: $M = 357$ Hz; /u/: $M = 370$ Hz; /ũ/: $M = 390$ Hz), intermediate for the mid vowel /ø/ ($M = 471$ Hz), and high for the open vowel /a/ ($M = 646$ Hz). Statistical testing revealed that the effect for vowel identity was significant ($\chi^2(4) = 3,927.9$, $p < .001$). The effects for stress, speaker gender and phonemic length as well as the interactions involving these factors were not significant (all $ps > .05$). Pairwise tests indicated that the /i ~ u/ contrast was not significant ($p > .1$). All other vowel pairs showed significant differences (all $ps < .01$). This indicates that the close vowels /i u/ were produced higher than the /ũ/, the /i u u/ were all higher than the /ø/, and the /ø/ was in turn higher than the /a/ (i.e. /i u/ > /ũ/ > /ø/ > /a/).

For F2, the highest value was observed for the front vowel /i/ ($M = 2,323$ Hz), followed by the /ũ/ ($M = 1,564$ Hz), the /a/ ($M = 1,548$ Hz), the /u/ ($M = 1,271$ Hz), and the /ø/ ($M = 1,263$ Hz). Mixed modelling showed significant effects for vowel identity ($\chi^2(4) = 2,781.9$, $p < .001$) and lexical stress ($\chi^2(1) = 7.9$, $p = .005$) as well as a significant interaction between

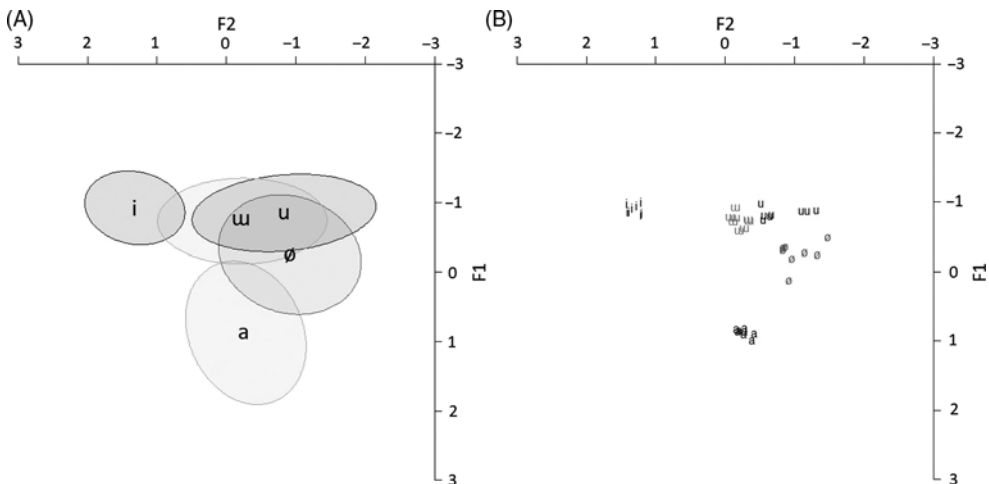


Figure 2 F1 \times F2 vowel space for the five phonemes /i ũ u ø a/ identified in Givón (2016); z-score-transformed group means (A) and individual speaker means (B); ellipses correspond to \pm one standard deviation.

pausing before/after the target word in the carrier sentence, etc.). As such, questions related to the differences between citation forms and connected speech must remain outside the scope of the current study and would need to be addressed in future research using a different set of data.

stress and length ($\chi^2(1) = 5.5, p = .02$). The effects for length and gender and the rest of the interactions did not reach significance (all $ps > .1$). Pairwise comparisons showed that the differences in F2 were not significant in two pairs: /u ~ a/ and /u ~ ø/ (both $ps > .05$). For all other pairs, differences in F2 were significant (all $ps < .001$). This shows that the /i/ was front, the /u/ and the /a/ were central, and the /u/ and the /ø/ were back (i.e. /i/ > /u a/ > /u ø/). Furthermore, long unstressed vowels had higher F2 than long stressed vowels ($M_{Unstressed} = 1,652$ Hz, $M_{Stressed} = 1,611$ Hz; $p = .04$), which signals a slightly more retracted position for the long stressed vowel category in general.

For the allophonic patterns, the first alternation involves the /ø/ that is expected to retract to [ɔ] when adjacent to a velar or a uvular consonant (Givón 2016). For F1, statistical testing showed a significant fixed effect for the following environment ($\chi^2(1) = 6.2, p = .01$), with vowels showing higher mean F1 when followed by back consonants ($M = 481$ Hz) than other sounds ($M = 465$ Hz). The following environment was also significant for F2 ($\chi^2(1) = 4.4, p = .04$), with lower values observed in front of velars and uvulars ($M = 1,235$ Hz) than elsewhere ($M = 1,350$ Hz). Effects and interactions involving the preceding environment and stress were not significant for either F1 or F2 (all $ps > .1$). (The interaction of phonemic length and environment could not be investigated statistically due to the limited number of long /ø/ tokens.) Taken together, the findings for F1 and F2 indicate only slight lowering and retraction of the vowel in front of back consonants. This can be seen in Figure 3, which shows z-score-transformed F1 and F2 values for the prototypical realization of the /ø/ and its positional allophone.

For the low central vowel /a/, Givón (2016) suggests that it undergoes either fronting to [æ] or both fronting and raising to [ɛ] when it is either preceded or followed by /j i ø/. For F1, statistical modeling showed no significant effects or interactions involving environment, stress or length (all $ps > .05$). For F2, there was a significant effect for the preceding environment ($\chi^2(1) = 12, p < .001$), with F2 being slightly higher when the /a/ was preceded by /j i ø/ ($M = 1,577$ Hz) than when preceded by other sounds ($M = 1,541$ Hz). The rest of the main effects and interactions did not reach significance for F2 (all $ps > .1$). This indicates slight fronting but not raising of the vowel, which can be seen in Figure 4. The figure shows z-score-transformed group means (panel A) and individual speaker means (panel B) for the prototypical realization of the /a/ and its positional allophone.

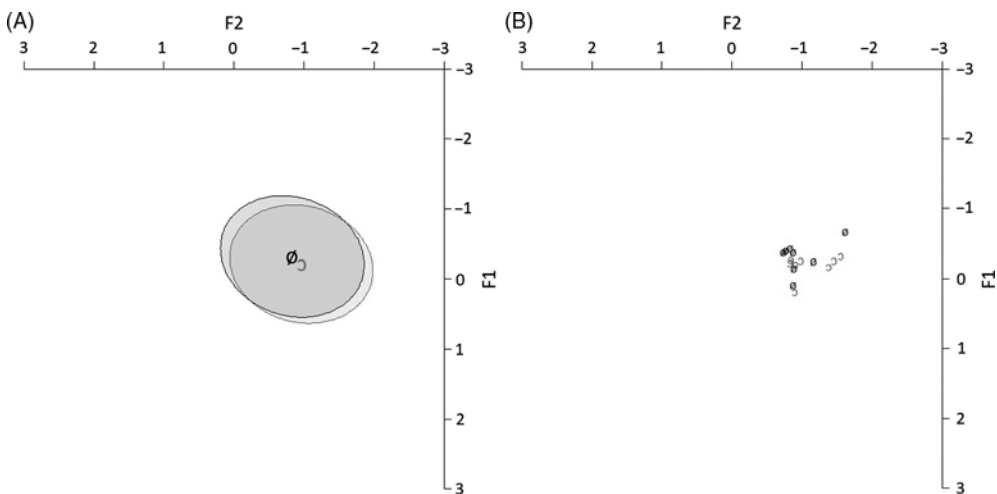


Figure 3 F1 × F2 vowel space for the /ø/ ~ [ɔ] alternation; z-score-transformed group means (A) and individual speaker means (B) for the prototypical realization (marked as ø; in black) and the positional allophone (marked as ɔ; in grey); ellipses correspond to +/- one standard deviation.

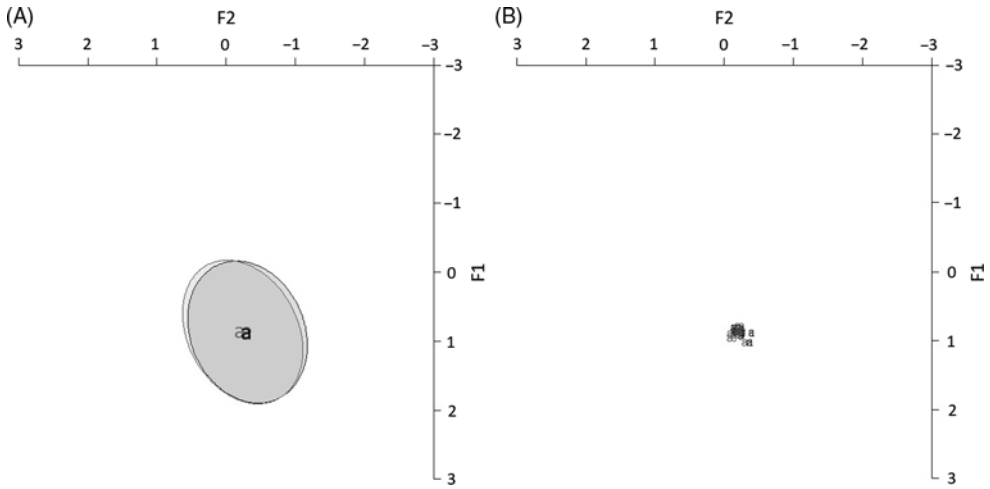


Figure 4 F1 × F2 vowel space for the /a/ ~ [ɛ/æ] alternation; z-score-transformed group means (A) and individual speaker means (B) for the prototypical realization (marked as a; in black) and the positional allophone (marked as æ; in grey); ellipses correspond to +/- one standard deviation.

The /u/ ~ [i] alternation is described in Givón (2011) as optional fronting of the /u/ in unstressed syllables after labial, dental, and palatal consonants. For F1, the unstressed subset of the /u/ tokens showed no significant effects for either the preceding or following environment (both $ps > .1$). For F2, there was a significant effect for the preceding environment ($\chi^2(1) = 5.1, p = .02$). F2 was lower after front consonants ($M = 1,540$ Hz) compared to other sounds ($M = 1,632$ Hz). The effect for the following environment was not significant for F2 ($p > .1$). (The small number of long /u/ tokens made it impossible to investigate the length × environment interaction.) As reflected in Figure 5, which shows z-score-transformed formant values for the prototypical realization of the /u/ and the positional allophone, this indicates retraction rather than fronting of the vowel after labial, dental, and palatal consonants.

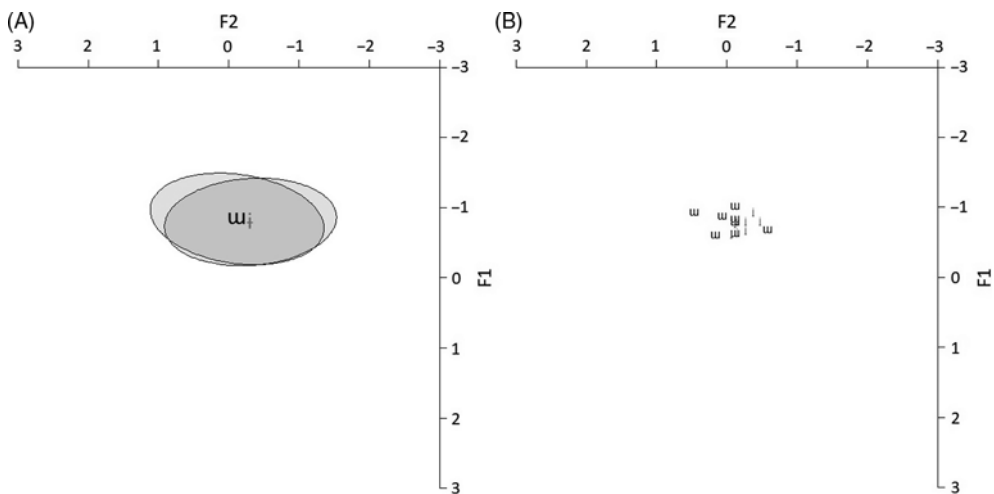


Figure 5 F1 × F2 vowel space for the /u/ ~ [i] alternation; z-score-transformed group means (A) and individual speaker means (B) for the prototypical realization (marked as u; in black) and the positional allophone (marked as i; in grey); ellipses correspond to +/- one standard deviation.

3.2 Vowel duration

For vowel duration, mixed modelling showed significant effects for phonemic length ($\chi^2(1)=103.1, p < .001$), stress ($\chi^2(1) = 26.3, p < .001$), vowel identity ($\chi^2(4) = 33.4, p < .001$), speaker gender ($\chi^2(1) = 21.2, p < .001$), and word-finality ($\chi^2(1) = 191.7, p < .001$). There were also significant interactions between length and vowel identity ($\chi^2(4) = 14, p = .007$) and stress and vowel identity ($\chi^2(4) = 42.6, p < .001$). The interactions involving stress, length, gender and word-finality were not significant (all $ps > .05$).

As reflected in Figure 6, phonologically long vowels showed greater duration than phonologically short vowels ($M_{Short} = 113$ ms, $SD = 54$; $M_{Long} = 173$ ms, $SD = 68$). In pairwise comparisons, the difference was significant for all five vowel pairs (/i/: $M = 119$ ms, $SD = 57$; /i:/: $M = 152$ ms, $SD = 60$; /u/: $M = 103$ ms, $SD = 61$; /u:/: $M = 193$ ms, $SD = 86$; /u/: $M = 103$ ms, $SD = 51$; /u:/: $M = 153$ ms, $SD = 58$; /ø/: $M = 111$ ms, $SD = 55$; /ø:/: $M = 193$ ms, $SD = 84$; /a/: $M = 113$ ms, $SD = 51$; /a:/: $M = 181$ ms, $SD = 67$; all $ps < .001$).

The interaction between stress and vowel identity can be seen in Figure 7. Unstressed vowels were shorter compared to the stressed realizations ($M_{Unstressed} = 112$ ms, $SD = 55$, $M_{Stressed} = 137$ ms, $SD = 66$). Pairwise testing indicated that the difference was not significant for the /a/ ($M_{Unstressed} = 115$ ms, $SD = 52$, $M_{Stressed} = 140$ ms, $SD = 68$; $p > .05$) but it was significant for all other vowels (/i/: $M_{Unstressed} = 119$ ms, $SD = 58$, $M_{Stressed} = 131$ ms, $SD = 57$; /u/: $M_{Unstressed} = 98$ ms, $SD = 58$, $M_{Stressed} = 143$ ms, $SD = 80$; /u/: $M_{Unstressed} = 104$ ms, $SD = 54$, $M_{Stressed} = 130$ ms, $SD = 57$; /ø/: $M_{Unstressed} = 101$ ms, $SD = 53$, $M_{Stressed} = 137$ ms, $SD = 65$; all $ps < .05$).

Vowels were also longer in word-final syllables than in non-final positions ($M_{Word-final} = 169$ ms, $SD = 72$; $M_{Non-final} = 114$ ms, $SD = 55$). This was independent of phonemic length, stress or gender. Vowels also had slightly longer duration when produced by female than male speakers ($M_{Female} = 123$ ms, $SD = 62$; $M_{Male} = 117$ ms, $SD = 56$), which was also independent of other factors.

3.3 Spectral emphasis

For spectral emphasis, statistical testing revealed significant main effects for stress ($\chi^2(1) = 166.7, p < .001$), vowel identity ($\chi^2(4) = 466.2, p < .001$), and speaker gender ($\chi^2(1) = 1,277.6, p < .001$). There were also significant interactions between stress and vowel identity

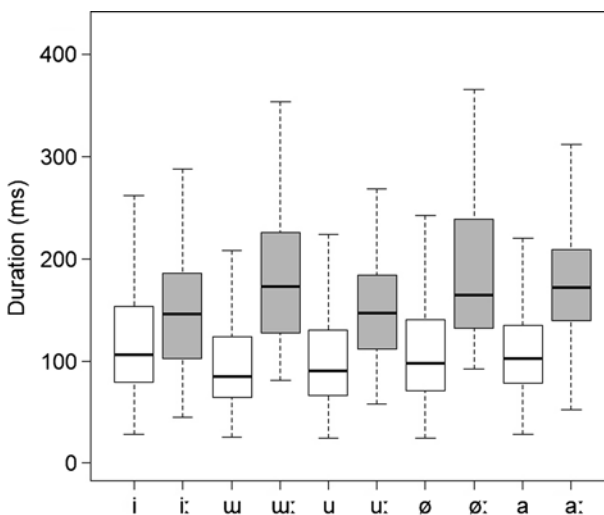


Figure 6 Duration (ms) for underlyingly short (unshaded) vs. long (shaded) vowels.

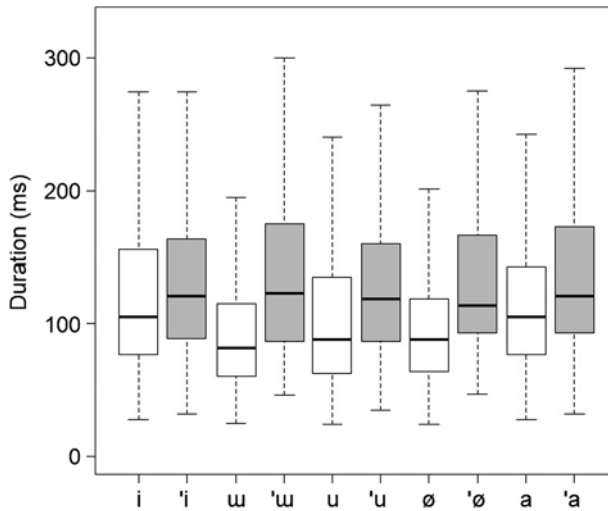


Figure 7 Duration (ms) for the unstressed (unshaded) vs. stressed (shaded) realizations of the five vowels.

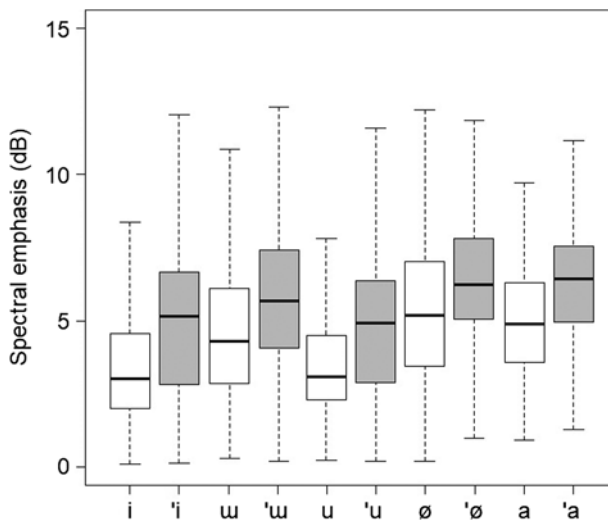


Figure 8 Spectral emphasis (dB) for the unstressed (unshaded) vs. stressed (shaded) realizations of the five vowels.

($\chi^2(4) = 20.7, p < .001$) and stress and speaker gender ($\chi^2(1) = 97.3, p < .001$). Phonemic length did not reach statistical significance and did not interact with stress (both $ps > .1$).

Figure 8 provides the means for the vowel phonemes separated by stress. Stressed realizations always showed greater spectral emphasis than their unstressed counterparts ($M_{Unstressed} = 4$ dB, $SD = 2.31$; $M_{Stressed} = 5.4$ dB, $SD = 2.9$). The increase in spectral emphasis was significant in pairwise tests for all five phonemes (/i/: $M_{Unstressed} = 3.2$ dB, $SD = 1.8$, $M_{Stressed} = 4.8$ dB, $SD = 2.8$; /u/: $M_{Unstressed} = 4.2$ dB, $SD = 2$, $M_{Stressed} = 5.6$ dB, $SD = 2.5$; /ø/: $M_{Unstressed} = 3.3$ dB, $SD = 1.9$, $M_{Stressed} = 4.9$ dB, $SD = 2.8$; /a/: $M_{Unstressed} = 4.4$ dB, $SD = 2.3$, $M_{Stressed} = 5.7$ dB, $SD = 3.1$; /a/: $M_{Unstressed} = 5.3$ dB, $SD = 2.6$, $M_{Stressed} = 6.4$ dB, $SD = 2.8$; all $ps < .001$). These results indicate that the five vowels were all produced louder when stressed, with higher spectral emphasis values seen for the more open vowels.

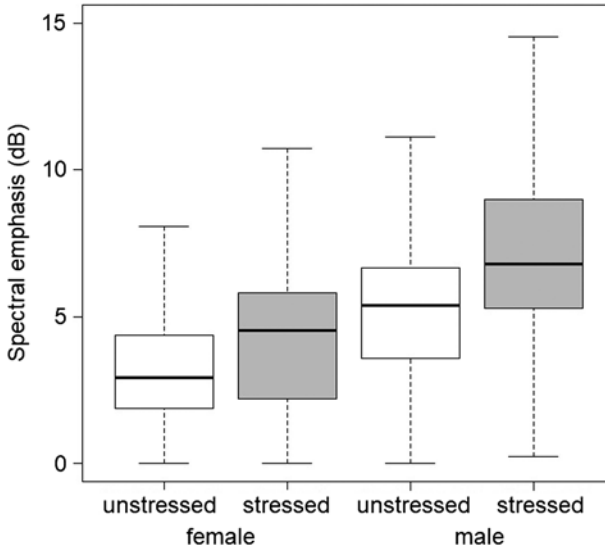


Figure 9 Spectral emphasis (dB) for the unstressed (unshaded) and stressed (shaded) vowels produced by female vs. male speakers.

The interaction between stress and speaker gender is shown in **Figure 9**. The figure provides spectral emphasis values for the unstressed and stressed vowels produced by the female speakers ($M_{Unstressed} = 3.4$ dB, $SD = 2.2$; $M_{Stressed} = 4.2$ dB, $SD = 2.4$) and the male speakers ($M_{Unstressed} = 5.3$ dB, $SD = 2$; $M_{Stressed} = 7.3$ dB, $SD = 2.6$). Pairwise comparisons revealed that the differences between unstressed and stressed vowels were significant both within and across genders (all $ps < .001$). This suggests that the stressed vowels were produced with greater loudness by both genders and that male speakers showed higher overall spectral emphasis as well as a greater increase in spectral emphasis from the unstressed to stressed categories.

3.4 f₀

For f₀, statistical modelling showed significant effects for stress ($\chi^2(1) = 373.4$, $p < .001$), vowel identity ($\chi^2(4) = 83$, $p < .001$), length ($\chi^2(1) = 12.6$, $p < .001$), and speaker gender ($\chi^2(1) = 2,459.4$, $p < .001$). There were also significant interactions between stress and vowel identity ($\chi^2(4) = 17.88$, $p = .001$) and stress and length ($\chi^2(1) = 17.5$, $p < .001$). The interaction between gender and stress was not significant ($p > .1$).

Figure 10 provides the results for the five phonemes separated by stress. As can be seen in the figure, stressed vowels showed higher f₀ than unstressed vowels ($M_{Unstressed} = 162$ Hz, $SD = 33$; $M_{Stressed} = 180$ Hz, $SD = 35$). Pairwise tests indicated that the difference between stressed and unstressed realizations was significant for all 5 phonemes (/i/: $M_{Unstressed} = 162$ Hz, $SD = 35$, $M_{Stressed} = 182$ Hz, $SD = 35$; /u/: $M_{Unstressed} = 168$ Hz, $SD = 33$, $M_{Stressed} = 179$ Hz, $SD = 38$; /u/: $M_{Unstressed} = 165$ Hz, $SD = 33$, $M_{Stressed} = 188$ Hz, $SD = 35$; /ø/: $M_{Unstressed} = 167$ Hz, $SD = 28$, $M_{Stressed} = 189$ Hz, $SD = 31$; /a/: $M_{Unstressed} = 159$ Hz, $SD = 32$, $M_{Stressed} = 176$ Hz, $SD = 35$; all $ps < .001$).

The interaction between stress and length is illustrated in **Figure 11**, which shows f₀ for the unstressed versus stressed short and long vowels (short: $M_{Unstressed} = 163$ Hz, $SD = 33$, $M_{Stressed} = 182$ Hz, $SD = 33$; long: $M_{Unstressed} = 147$ Hz, $SD = 30$, $M_{Stressed} = 176$ Hz, $SD = 39$). In pairwise tests, only the difference between short and long stressed vowels was not significant ($p > .1$). All other pairs showed significant differences (all $ps < .05$). This indicates that

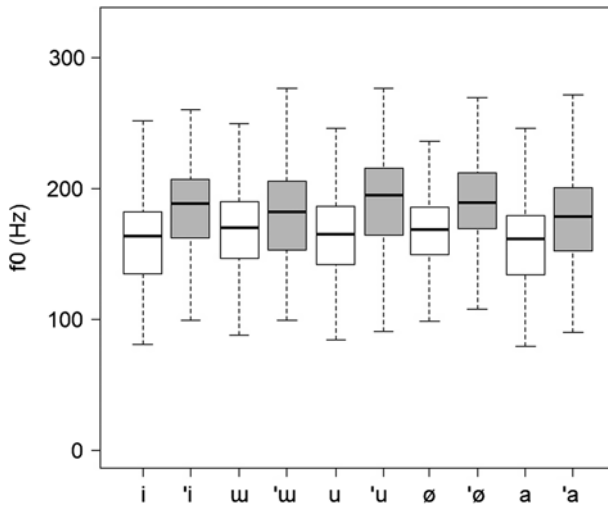


Figure 10 f0 (Hz) for the unstressed (unshaded) vs. stressed (shaded) realizations of the five vowels.

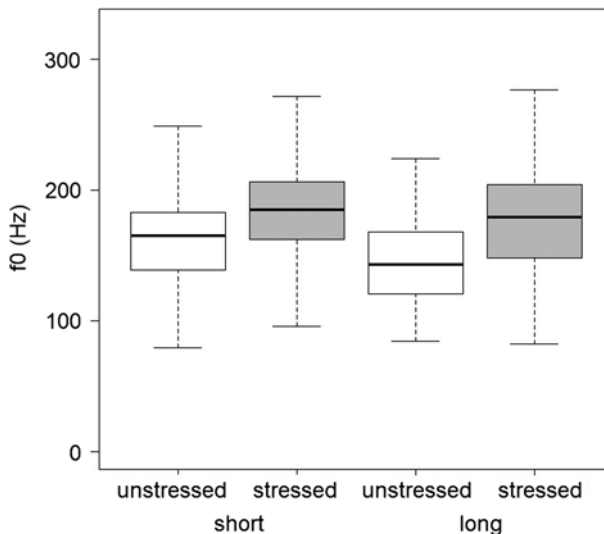


Figure 11 f0 (Hz) for the unstressed (unshaded) vs. stressed (shaded) short and long vowels.

both short and long vowels had higher f0 when stressed and that the lowest f0 was seen for the unstressed long vowels.

For the effect of speaker gender, female speakers showed higher f0 than male speakers ($M_{Female} = 184$ Hz, $SD = 27$; $M_{Male} = 140$ Hz, $SD = 28$). In the absence of an interaction between gender and stress, this reflects a higher overall pitch in female speech.

4 Discussion

In this study, we investigated the phonetic realization of Southern Ute vowels, including the general vowel space, vowel allophony, vowel quantity, and stress. Our goal was to provide instrumental phonetic data that can be compared to the previous auditory accounts of the language and can be used as a baseline for future research, which will facilitate language

revitalization efforts. The findings are based on recordings from eight fluent speakers of Southern Ute from the last generation to acquire the language as their L1. We analyzed over 6000 vowel tokens, which made it possible to not only describe the observed patterns but also model them statistically.

Our results for F1 and F2 show five distinct vowel categories, including three close vowels, one mid vowel, and one open vowel. The number of categories and their overall distribution across the three height levels are consistent with the five-phoneme inventory in Givón (2016) that also includes three close, one mid, and one open vowel. The size of the inventory is also in line with other Numic languages (e.g. Southern and Northern Paiute, Sapir 1930, Thornes 2003) as well as the cross-linguistic preference for a five-phoneme vowel system (Schwartz et al. 1997). Furthermore, the vowels /i u a/ match Givón's descriptions with respect to backness, with /i/ being front, /u/ being back, and /a/ being central. F2 data also reveal fronting of long unstressed vowels. This has not been mentioned in the previous accounts of Southern Ute but has been attested in other languages with lexical stress that also show an interaction between stress and backness (e.g. less peripheral articulation of unstressed vowels; Delattre 1969, Ortega-Llebaria & Prieto 2011).

Some of our findings for vowel formants do not fully support the previous accounts. For example, the /u/ is described in Givón (2011, 2016) as a close back unrounded vowel. Our data indicate that its F2 is significantly higher than the F2 of the back rounded /u/ and not significantly different from the F2 of the open central vowel /a/. This suggests that the vowel is more central than previously described. It also has significantly lower F1 than either /i/ or /u/ but higher F1 than the /ø/, which suggests a near-close placement. Thus, the phoneme would be better represented as the close central vowel /i/, with its phonetic realization being the near-close centralized [i̠], or as the close-mid central vowel /ø/, with a slightly raised phonetic realization. This would be in line with the recent instrumental accounts of the corresponding vowel phoneme being either /i/ or /ø/ in related Numic languages, including Comanche, Kawaiisu and Northern Paiute (Haynes 2010, Herrick 2011, Thomas 2017). It would also be consistent with the general preference for avoiding back unrounded vowels that is seen in the world's languages (Lindblom 1986, Schwartz et al. 1997).

Givón (2011) also specified that the /u/ was fronted in unstressed syllables after labial, dental and palatal consonants. This pattern was omitted without comment from the updated chapter on Southern Ute sounds in Givón (2016), suggesting that the alternation may have disappeared due to sound change or it may have been misidentified. Our results show that instead of triggering fronting, adjacency to non-back consonants leads to slight backing of the vowel. This type of dissimilatory behavior has not been described in any of the previous auditory accounts of Southern Ute, and it is also rare and highly constrained cross-linguistically (see Flemming 2003). One possible explanation that will need to be investigated in future research is that if the vowel was originally a back /u/ that has undergone centralization, speakers may be using backing as means of maintaining articulatory distinctness, since the centralization of /u/ makes it more distinct from the back consonants but less distinct from the front consonants.

Results for the /ø/ are also not fully in line with the earlier descriptions that treat the /ø/ as a mid front rounded vowel that gets retracted to [ɔ̠] next to velar and uvular consonants (Givón 2011, 2016). Determining the status of the /ø/ is especially important considering that the inventories of related Numic languages contain either the close-mid back rounded /o/ or the open-mid back rounded /ɔ/ rather than the front rounded /ø/ (e.g. Northern and Southern Paiute, Panamint, Shoshoni, Comanche, Proto Uto-Aztecan; Sapir 1930, Langacker 1970, Charney 1993, Elzinga 1999, Gould & Loether 2002, Thornes 2003, McLaughlin 2006, Haynes 2010). A vowel inventory containing the /ø/ rather than the /o/ (or /ɔ/) would also be rare cross-linguistically (Maddieson 2013) and would not be maximized according to the Dispersion Theory (Liljencrants & Lindblom 1972, Becker-Kristal 2010). Our data indicate that when looking at the phoneme in general (i.e. by averaging across the two allophones), it does not differ significantly in F2 from the /u/. Furthermore, despite being reflected in both

orthographies, only slight retraction and slight lowering happen in front of back consonants. Even at the level of individual tokens, the vast majority of productions are either central or back. This shows that the prototypical realization of the vowel is substantially further back than previously reported. As there are no recordings to accompany Givón's original field notes, it is not possible to determine whether the vowel was always non-front or whether it underwent sound change that led to retraction. However, the current findings clearly indicate that in present day Southern Ute, the phoneme would be more accurately represented either as an /ø/ that undergoes retraction to [o] (or [ɔ]) next to back consonants or an /o/ (or /ɔ/) that gets fronted to [ə] next to non-back sounds. The latter analysis would also bring the vowel inventory of Southern Ute in line with the inventories of other Numic languages as well as the cross-linguistic preference for the back rounded vowels.

For the /a/ ~ [ɛ/æ] alternation, our study offers only partial support for the previous accounts in Givón (2011, 2016). The allophony is supposed to be conditioned by the presence of /j i ø/ either before or after the vowel, with the positional allophone being [æ] for older speakers and [ɛ] for younger speakers (who would have been adults in their 30s through 50s as there were no more children or young adults speaking the language fluently in the 1970s and 1980s when Givón's data were collected). Our results show that only the preceding environment plays a role at a statistically significant level. Furthermore, although some of the individual tokens clearly fall within the [ɛ]-range, at the group level this allophony involves slight fronting but not raising of the vowel. This suggests that the most common positional allophone is [æ]. Considering that Givón's data were gathered over 40 years ago, the speakers who took part in our study are the original 'younger' group and, if the previous description of the age-based difference is accurate, they appear to have transitioned from [ɛ] to [æ].

The findings for segmental duration substantially expand our understanding of how the phonemic length contrast is realized in Southern Ute. As noted in the Introduction, there is no phonetic description for this contrast in Givón (2016), while Charney (1996) states that long vowels are two times longer than short vowels. This is also how long vowels are often described in the grammars and dictionaries of other Uto-Aztecan languages (e.g. for the long vowels of Shoshoni, 'pronunciation is generally held twice as long as for a short vowel', Gould & Loether 2002: 10). Our data confirm that phonemically long vowel have greater duration, which is consistent with the findings for other languages with phonemic vowel length differences (e.g. Chickasaw – Gordon 2004; Yakut – Vasilyeva, Arnhold & Järviö 2016). At the same time, durational differences are fairly modest at the group level, with the long vowels not being twice as long as the short vowels (as suggested in Charney 1996). Vowel lengthening is also seen in final positions, which has not been mentioned in the previous accounts of Southern Ute despite being attested in other Uto-Aztecan languages (e.g. Northern Paiute; Haynes 2010) as well as many other language families (e.g. the Muskogean family – Johnson & Martin 2001, Gordon & Munro 2007; the Indo-European family – Delattre 1966).

We also observed a significant relationship between vowel length and stress. The use of duration as a cue to stress has long been established in the literature (Fry 1955, Lieberman 1960, Klatt 1976). However, since vowel length is already contrastive in Southern Ute, it was important to determine whether speakers rely on duration to signal not only vowel quantity but also stress (see Berinstein 1979). Our results reveal modest but statistically significant lengthening of the majority of vowels in stressed positions, which confirms that Southern Ute speakers use duration to differentiate between stressed and unstressed vowels. The observed differences are more limited than what is often found in languages without phonemic length differences (e.g. English; Lunden 2017) but they are in line with the findings for languages with a phonological vowel quantity contrast (e.g. Chickasaw; Gordon 2004). Consequently, the claim in Givón (2016) that duration and stress do not interact may hold true at the phonological level but not necessarily at the phonetic level.

Results for spectral emphasis provide further insight into the phonetics of stress in Southern Ute. As noted in the Introduction, spectral emphasis is a measure of vocal effort.

Sounds produced with greater vocal effort show increased intensity and more high-frequency content (Allen 1971, Ternstrom, Bohman & Sodersten 2006). As stressed vowels have greater intensity than their unstressed counterparts and these differences are largely limited to the upper parts of the spectrum (Sluijter & van Heuven 1996), spectral emphasis measures have been found to be a reliable cue to stress in both production and perception of speech (Sluijter & van Heuven 1996, Sluijter, van Heuven & Pacilly 1997, Haynes 2003, Eriksson et al. 2016). In our study, we observed higher spectral emphasis values for the stressed category, which signals that stressed vowels were louder than their unstressed counterparts. This matches the results for spectral emphasis in other languages (e.g. English, Italian, Swedish; Haynes 2003, Eriksson & Heldner 2015, Eriksson et al. 2016) and, more generally, the finding that higher intensity is associated with the stressed category (e.g. in Chickasaw; Gordon 2004). Spectral emphasis was also higher for the more open vowels, which is in line with previous research that has shown an inverse relationship between vowel height and loudness (e.g. Lehiste 1970). Furthermore, we found higher overall levels of spectral emphasis and a greater increase in values (from unstressed to stressed) in male speakers. This is consistent with the findings for other languages that also show greater loudness and a more prominent difference between unstressed and stressed vowels in male speakers (e.g. English and Swedish; Traunmuller & Eriksson 2000, Eriksson & Heldner 2015, Eriksson et al. 2016).

The fifth and final acoustic measure investigated in the current study was f_0 . Fundamental frequency is a correlate of pitch accent (Huss 1978, Ladd 1996, Sluijter & van Heuven 1996, Gordon 2014). In languages with lexical stress, pitch accents are usually realized on stressed syllables in words under focus, with stressed vowels showing higher f_0 when accented (e.g. see the English findings in Sluijter & van Heuven 1996). There are, however, stress languages that do not associate pitch accents with stressed syllables (e.g. Kuot; Lindstrom & Remijsen 2005), so it was important to determine the relationship between accent and stress in Southern Ute. Given the nature of our dataset (i.e. words produced in either isolation or embedded in the same declarative carrier sentence), all word tokens were produced under focus and with declarative intonation, which is when stress and accent can be expected to covary. Our results showed higher f_0 for the stressed vowels, which confirms that stress is cued by pitch accent in declaratives in Southern Ute. There was also an interaction with phonemic length, with the lowest f_0 observed for the long unstressed vowels. This is consistent with Vasilyeva, Arnhold & Järvikivi (2016) who found lower f_0 for long Yakut vowels, presumably because they are long enough to allow for changes to the fundamental frequency. We also observed higher f_0 values in female tokens, which is the usual finding given the physiological differences in vocal tract morphology between female and male speakers (e.g. Hillenbrand et al. 1995).

As our study only aimed to provide an initial description of the phonetic realization of Southern Ute vowels, we leave a number of issues for future research. One such issue is the source of the substantial variability that is seen within each of the five phonemic categories. While it is not uncommon for languages with smaller inventories to allow for substantial allophony (e.g. Kabardian; Gordon & Applebaum 2006), only three allophonic patterns have so far been identified for Southern Ute. Determining the full system of allophonic relations affecting the vowels also requires examination of additional parameters that speakers may be relying on, including the role of such factors as rhoticity that can be used to enhance distinctness of allophones. It is also important to investigate the extent to which such patterns are phonologized (as presented in Givón 2011, 2016) or whether they are purely phonetic coarticulatory effects that may not be perceptually salient (as suggested by the modest differences in acoustic measures in our study). Our findings for f_0 also provide only an initial insight into the intonational system of Southern Ute. Future research will need to examine pitch accent in words produced without intonational focus as well as to provide a general description of the intonational patterns of the language. Additional research is also needed to evaluate how the pronunciation of vowels changes in connected speech (see Strange et al. 2007) and how the productions of fluent speakers who acquired the language as L1 differ from the speech of non-fluent speakers who learned Southern Ute as L2 (see Haynes 2010).

5 Conclusions

This study investigated the phonetic realization of vowels in Southern Ute, a highly endangered and underdescribed Native American language. The observed phonetic values match some of the earlier auditory descriptions of the language, such as the overall number of phonemic categories, the three-level height system, and the backness of three (out of five) vowels. Other previous accounts find either partial or no support, including tongue backness of the remaining two vowel phonemes as well as the descriptions of allophonic processes affecting vowels. We also show how stress and vowel quantity are cued in the language. These results help us understand how the vowel system of Southern Ute relates to the systems of other Uto-Aztec languages and the literature on vowels in general. The current findings also provide a quantitative baseline for future research and revitalization efforts.

Acknowledgments

We express our deepest gratitude to the Southern Ute speakers who took part in the study. We are also very grateful for the feedback received from the editor, anonymous reviewers, and audiences at scholarly meetings.

References

- Allen, George D. 1971. Acoustic level and vocal effort as cues for the loudness of speech. *Journal of the Acoustical Society of America* 49, 1831–1841.
- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255–278.
- Becker-Kristal, Roy. 2010. *Acoustic typology of vowel inventories and Dispersion Theory: Insights from a large cross-linguistic corpus*. Ph.D. dissertation, UCLA.
- Beckman, Mary. 1986. *Stress and non-stress accent*. Dordrecht: Foris.
- Berinstein, Ava E. 1979. A cross-linguistic study on the perception and production of stress. *UCLA Working Papers in Phonetics* 47, 1–59.
- Boersma, Paul & David Weenink. 2017. Praat: Doing phonetics by computer (computer program); version 6.0.28. <http://www.praat.org/> (accessed 20 March 2017).
- Charney, Jean. 1993. *A grammar of Comanche*. Lincoln, NE: University of Nebraska Press.
- Charney, Jean. 1996. *Ute dictionary*. Ignacio, CO: Southern Ute Tribe.
- Chladkova, Katerina, Paola Escudero & Paul Boersma. 2011. Context-specific acoustic differences between Peruvian and Iberian Spanish vowels. *The Journal of the Acoustical Society of America* 130, 416–428.
- Delattre, Pierre. 1966. A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics* 4, 183–198.
- Delattre, Pierre. 1969. An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics* 7, 295–325.
- Elzinga, Dirk. 1999. *The consonants of Gosiute*. Ph.D. dissertation, University of Arizona.
- Eriksson, Anders, Pier M. Bertinetto, Mattias Heldner, Rosalba Nodari & Giovanna Lenoci. 2016. The acoustics of lexical stress in Italian as a function of stress level and speaking style. *Proceedings of Interspeech 2016, San Francisco, CA, September 8–12, 2016*, 1059–1063.
- Eriksson, Anders & Mattias Heldner. 2015. The acoustics of word stress in English as a function of stress level and speaking style. *Proceedings of Interspeech 2015, Dresden, Germany, September 6–10, 2015*, 41–45.
- Flemming, Edward. 2003. The relationship between coronal place and vowel backness. *Phonology* 20, 335–373.
- Fry, Denis B. 1955. Duration and intensity as acoustic correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27, 765–768.

- Gay, Thomas. 1978. Physiological and acoustic correlates of perceived stress. *Language and Speech* 21, 347–353.
- Givón, Talmy. 1979. *Ute dictionary*. Ignacio, CO: Southern Ute Tribe.
- Givón, Talmy. 1980. *Ute reference grammar*. Ignacio, CO: Southern Ute Tribe.
- Givón, T[almy]. 2011. *Ute reference grammar*. Philadelphia, PA: John Benjamins.
- Givón, T[almy]. 2016. *Ute dictionary*. Philadelphia, PA: John Benjamins.
- Glasberg, Brian & Brian Moore. 1990. Derivation of auditory filter shapes from notched noise data. *Hearing Research* 47, 103–137.
- Gordon, Matthew. 2004. A phonological and phonetic study of word-level stress in Chickasaw. *International Journal of American Linguistics* 70, 1–32.
- Gordon, Matthew. 2014. Disentangling stress and pitch accent: Toward a typology of prominence at different prosodic levels. In Harry van der Hulst (ed.), *Word stress: Theoretical and typological issues*, 83–118. Cambridge: Cambridge University Press.
- Gordon, Matthew & Ayla Applebaum. 2006. Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association* 36, 159–186.
- Gordon, Matthew & Pamela Munro. 2007. A phonetic study of final vowel lengthening in Chickasaw. *International Journal of American Linguistics* 73, 293–330.
- Goss, James A. 1961. *A short dictionary of Southern Ute*. Ignacio, CO: Southern Ute Tribe.
- Gould, Drusilla & Christopher Loether. 2002. *An introduction to the Shoshoni language*. Salt Lake City, UT: The University of Utah Press.
- Haynes, Erin. 2010. *Phonetic and phonological acquisition in endangered languages learned by adults: A case study of Numu (Oregon Northern Paiute)*. Ph.D. dissertation, University of California, Berkeley.
- Heldner, Mattias. 2003. On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics* 31, 39–62.
- Herrick, Dylan. 2011. On Comanche's central mid vowel. *International Journal of American Linguistics* 77, 373–396.
- Hillenbrand, James, Laura A. Getty, Michael J. Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America* 97, 3099–3111.
- Huss, Volker. 1978. English word stress in the post-nuclear position. *Phonetica* 35, 86–105.
- Johnson, Keith & Jack Martin. 2001. Acoustic vowel reduction in Creek: Effects of distinctive length and position in the word. *Phonetica* 58, 81–102.
- Klatt, Dennis. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America* 59, 1208–1221.
- Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
- Langacker, Ronald W. 1970. The vowels of Proto Uto-Aztecan. *International Journal of American Linguistics* 36, 169–180.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
- Lenth, Russell. 2018. *emmeans*: Estimated marginal means, aka least-squares means. R package version 1.1.3.
- Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America* 32, 451–454.
- Liljencrants, Johan & Bjorn Lindblom. 1972. Numerical simulations of vowel quality systems: The role of perceptual contrasts. *Language* 48, 839–862.
- Lindblom, Bjorn. 1963. Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America* 35, 1773–1781.
- Lindblom, Bjorn. 1986. Phonetic universals in vowel systems. In John J. Ohala & Jeri J. Jaeger (eds.), *Experimental phonology*, 13–44. Orlando, FL: Academic Press.
- Lindstrom, Eva & Bert Remijsen. 2005. Aspects of prosody of Kuot, a language where intonation ignores stress. *Linguistics* 43, 839–870.
- Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different listeners. *The Journal of the Acoustical Society of America* 49, 606–608.

- Lunden, Anya. 2017. Duration, vowel quality, and the rhythmic pattern of English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8, 27.
- Maddieson, Ian. 2013. Front rounded vowels. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/>.
- McCloy, Daniel R. 2016. *phonR*: Tools for phoneticians and phonologists. R package (version 1.0-7).
- McLaughlin, John. 2006. *Timbisha (Panamint)*. Munich: Lincom Europa.
- Moore, Brian & Brian Glasberg. 1983. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *The Journal of the Acoustical Society of America* 74, 750–753.
- Oberly, Stacey. 2008. *Phonetic analysis of Southern Ute with a discussion of Southern Ute language policies and revitalization*. Ph.D. dissertation, University of Arizona.
- Oberly, Stacey & Viktor Kharlamov. 2015. The phonetic realization of devoiced vowels in the Southern Ute language. *Phonetica* 72, 1–19.
- Oberly, Stacey, Dedra White, Arlene Millich, Mary Inez Cloud, Lillian Seibel, Crystal Ivey & Lorelei Cloud. 2015. Southern Ute grassroots language revitalization. *Language Documentation and Conservation* 9, 324–343.
- Ortega-Llebaria, Marta & Pilar Prieto. 2011. Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech* 54, 73–97.
- R Core Team. 2017. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Sapir, Edward. 1930. Southern Paiute: A Shoshonean language. *Proceedings of the American Society of Arts and Sciences* 65, 1–296.
- Schwartz, Jean-Luc, Louis-Jean Boe, Nathalie Vallee & Christian Abry. 1997. Major trends in vowel system inventories. *Journal of Phonetics* 25, 233–253.
- Simons, Gary F. & Charles D. Fennig (eds.). 2017. *Ethnologue: Languages of the world*, 20th edn. Dallas, TX: SIL International. <http://www.ethnologue.com>.
- Singmann, Henrik, Ben Bolker, Jake Westfall & Frederik Aust. 2017. *afex*: Analysis of factorial experiments. R package (version 0.18–0).
- Sluijter, Agaath M. C. & Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America* 100, 2471–2485.
- Sluijter, Agaath M. C., Vincent J. van Heuven & Jos J. A. Pacilly. 1997. Spectral balance as a cue in the perception of linguistic stress. *The Journal of the Acoustical Society of America* 101, 503–513.
- Strange, Winifred, Andrea Weber, Erika S. Levy, Valeriy Shafiro, Miwako Hisagi & Kanae Nishi. 2007. Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *The Journal of the Acoustical Society of America* 122, 1111–1129.
- Ternstrom, Sten, Mikael Bohman & Maria Sodersten. 2006. Loud speech over noise: Some spectral attributes, with gender differences. *The Journal of the Acoustical Society of America* 119, 1648–1665.
- Thomas, Patrick. 2017. The central vowel of Kawaiisu. *International Journal of American Linguistics* 83, 539–559.
- Thornes, Tim. 2003. *A Northern Paiute grammar with texts*. Ph.D. dissertation, University of Oregon.
- Traunmuller, Hartmut & Anders Eriksson. 2000. Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America* 107, 3438–3451.
- Vasilyeva, Lena, Anja Arnhold & Juhani Järvikivi. 2016. Phonetic correlates of phonological vowel quantity in Yakut read and spontaneous speech. *The Journal of the Acoustical Society of America* 139, 2541–2550.