Title
Phonetic Description of a Three-way Stop Contrast in Northern Paiute

Journal
UC Berkeley PhonLab Annual Report, 6(6)

ISSN
2768-5047

Author
Kataoka, Reiko

Publication Date
2010

DOI
10.5070/P78sz0x940

Phonetic description of a three-way stop contrast in Northern Paiute*

Reiko Kataoka

Abstract
This paper presents the phonetic description of a three-way phonemic contrast in the medial stops (lenis, fortis, and voiced fortis stops) of a southern dialect of Northern Paiute. Qualitative and quantitative analysis of VOT, closure duration, and voice quality was performed on field recordings of a female speaker from the 1950s. The findings include that: 1) voiced fortis stops are realized phonetically as voiceless unaspirated stops; 2) the difference between fortis and voiced fortis and between voiced fortis and lenis in terms of VOT is subtle; 3) consonantal duration is a robust acoustic characteristic differentiating the three classes of stops; 4) lenis stops are characterized by a smooth VC transition, while fortis stops often exhibit aspiration at the VC juncture, and voiced fortis stops exhibit occasional glottalization at the VC juncture. These findings suggest that the three-way contrast is realized by combination of multiple phonetic properties, particularly the properties that occur at the vowel-consonant boundary rather than the consonantal release.

1. Introduction

Northern Paiute (NP) belongs to the Western Numic branch of the Uto-Aztecan language family and is divided into two main dialect groups: the northern group, Oregon Northern Paiute, and the southern group, Nevada Northern Paiute (Nichols 1974:4). Some of the southern dialects of Nevada Northern Paiute, known as Southern Nevada Northern Paiute (SNNP) (Nichols 1974), have a unique three-way contrast in the medial obstruent: 'fortis', 'lenis', and what has been called by Numic specialists the 'voiced fortis' series. Other dialects of NP and the neighboring Mono language have a two-way contrast in the corresponding obstruent sets (Liljeblad 1966, Nichols 1974, Norris 1986, Thornes 2003). Although the existence of this three-way contrast has long been recognized, there is little detailed study of its phonetic properties, a descriptive gap this paper fills by describing the important aspects of medial obstruent contrast in SNNP.

A unique methodological choice was made in the present study. As Maddieson and Ladefoged (1985: 435) point out, a large empirical study treats data taken from many speakers so that the results may be generalized to the underlying population, and they may be considered to reveal properties of a language as a whole rather than of a particular speaker's idiolect. In this study, however, I examine the speech of a single speaker. As described in detail in Section 4.1, the speech data was taken from field recordings of a female speaker made in the early 1950s. Archived materials offer valuable speech data for the study of endangered languages such as SNPP. Linguistic properties of endangered languages tend to change rapidly under the strong influence of a dominant language, and archived speech materials can reveal phonetic details which may no longer be present in current speech forms, or the archived data might represent a unique regional variety which has already been lost. As such, archived materials have the

* This paper was originally written in spring 2007 as a Qualifying Paper. The present manuscript reflects grammatical corrections and stylistic modifications of the original manuscript. My sincere thanks to Ronald Sprouse for editorial assistance.

potential to serve as a valuable, and sometimes the only, source that reveals diachronic and/or regional variations of an endangered language.

The paper is organized as follows. Section 2 provides an overview of Numic languages, with particular focus on their phonological features, and an overview of SNNP. Section 3 reviews previous descriptions of the medial stops of Northern and Southern dialects of NP. In Section 4 I describe the methodology of the acoustic study and provide the results of qualitative and quantitative analysis of SNNP medial stops, and an interim summary is offered in Section 5. In Section 6 I discuss the possible perceptual consequences of the observed phonetic characteristics of medial stops and relate them to the sound pattern of the other dialects (Section 6.1); this section also discusses the typological significance of SNNP fortis and voiced fortis stops, which share some crucial acoustic properties with stop systems in other languages (Section 6.2). Finally, in Section 7, I discuss the implications these results have for the diachronic development of SNNP fortis and voiced fortis stops.

2. Background – Northern Paiute and Numic consonant alternation

In this section, I present an overview of NP and related languages and describe the morphophonological alternations of medial consonants in these languages. The phonemes that surface as a result of alternations vary across the languages, and these alternations are generally attributed to Proto Numic features (Nichols 1974), with the exception of preaspiration, which is an innovation in Proto Central Numic (see Miller, et al. 2005). Medial consonant alternations are observed not only at morpheme boundaries but also morpheme-internally, suggesting that today's medial consonant alternation might be a remnant of earlier morpheme boundary processes.

2.1. Language

Northern Paiute languages, together with Mono, forms the Western Numic sub-group of the Numic branch of the Uto-Aztecan language family and are spoken in the area "extending from the middle Snake and Owyhee River drainages, east to the Deschutes River, south across the northwestern third of the Great Basin region of eastern Oregon, western Nevada, and parts of eastern California to the Mono Lake area" (Thornes 2003). Other languages closely related to Northern Paiute are: Panamint, Shoshone[1] and Comanche of Central Numic subgroup, and Kawaiisu, Chemehuevi, Southern Paiute, and Ute of Southern Numic subgroup. The geographical distributions of the languages are shown in Figure 1.

Northern Paiute is divided into two dialect groups along a north-south dimension with a boundary running from between the Surprise and Honey Lake valleys in the west and the Owyhee drainage in the east (Nichols 1974; Liljeblad 1966). Nichols (1974:4) uses the term Oregon Northern Paiute for the northern group and Nevada Northern Paiute for the southern group. Nevada Northern Paiute has an exceptional sub-dialect group, which is called in this paper Southern Nevada Northern Paiute (SNNP). SNNP is unique in that it has a three-way contrast among medial obstruents—'fortis', 'lenis', and what has been called by Numic specialists 'voiced fortis' series—while all other dialects in Western Numic languages have a two-way distinction in the corresponding obstruent sets (Liljeblad 1966, Nichols 1966, Norris

[1] Alternative spelling for Shoshone is Shoshoni and Panamint is also known as Koso (Nichols 1974: 6)

1986, Thornes 2003). The next section describes these typologically rare phonological systems of the Numic languages in detail.

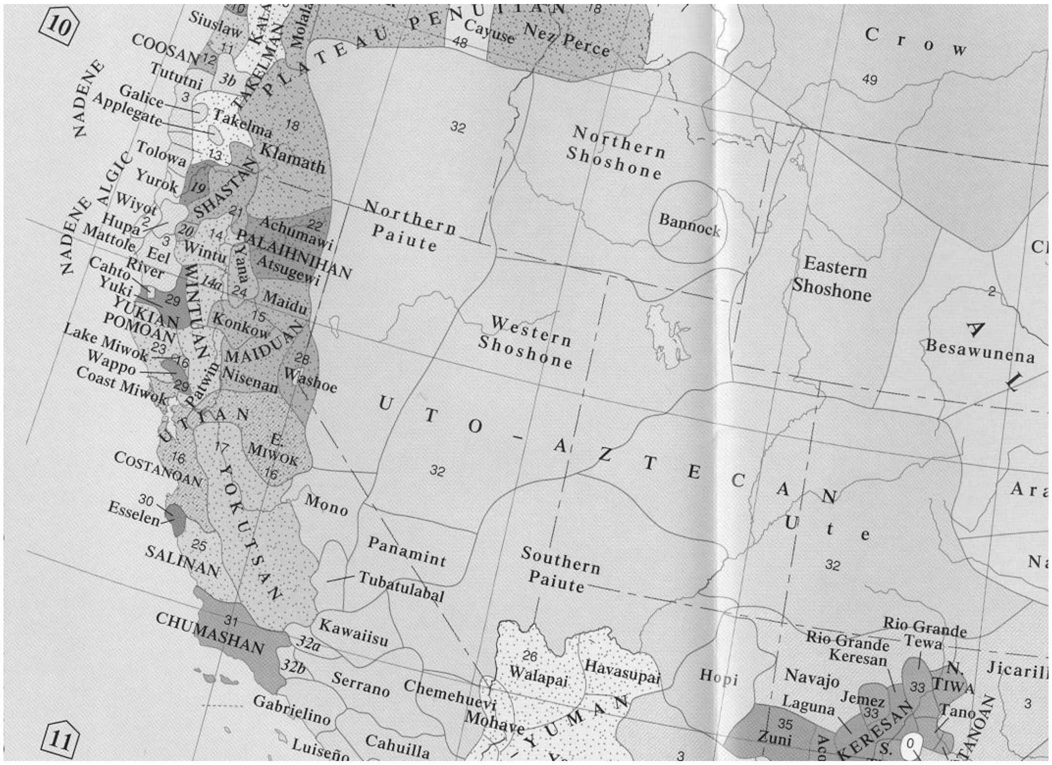


Fig.1. Numic Language Map (adopted from Goddard 1996)

2.2. Consonant alternation (Numic Final features)

Numic languages exhibit unique consonant alternation patterns in morphologically complex words, where the initial consonant of the second/third morpheme varies depending on the preceding morpheme. In traditional Numic studies, such alternations have been described in terms of 'final feature', or a lexically-specified feature of a morpheme, which, if another morpheme follows within a word, determines the manner of articulation of the initial consonant of the following morpheme (Sapir, 1930-31; Liljeblad, 1966; Nichols, 1974; Langacker, 1976; Miller, 1996; Thornes, 2003). The type of final features observed in the Numic languages are 'fortis', 'lenis', 'voiced fortis', 'prenasalization' and 'preaspiration', and the particular feature associated with a given morpheme varies among the languages. Numic specialists describe Proto Nunic morphophonemes with four final features. In this section, following Nichols, these features are referred to as 'series 1', 'series 2' 'series 3' and 'series 4', and their associated phonological alternation patterns are shown in Table 1.

The effects of final features are best illustrated by Shoshone examples because Shoshone preserves all four final feature distinctions, while all other Numic languages have lost one or more of these distinctions. Examples of Shoshone morphemes that are associated with each final feature are presented in Table 2, and their correspondences to following consonants are shown in

Table 3. For example, the initial stop in the word /pai/ 'to have' in its bare form is a plain voiceless stop /p/. The same stop will be realized as lenis stop /ß/, fortis stop /pp/, prenasalized stop /mb/, or voiceless fricative /ɸ/ when preceded by another morpheme /tsoo/ 'great grandparent', /tɨpa/ 'pinenuts', /tsoo/ 'beads', or /hai/ 'crow' respectively.

TABLE 1
Proto Numic final features and reflexes in current Numic languages (from Nichols 1974)

Final features	Effects in current Numic languages
*Series 1 (unmarked)	Lenition in all languages
*Series 2 (marked as ['])	Fortition in all languages that have series 2 feature
*Series 3 (marked as ["])	Penasalization in Shoshone and Southern Paiute Voiced fortis in SNNP Merged to series 2 in all other languages
*Series 4 (marked as [h])	(Pre)aspiration in all languages that have series 4 feature * This feature appears only before sonorant—hj, hw

TABLE 2
Shoshone morphemes with final features (from Miller 1996)

Series 1 (lenis)		Series 2 (fortis)		Series 3 (prenasal)		Series 4 (aspirate)	
nɨ	'my'	*tɨpa'*	'pinenuts'	*ɨ"*	'your'	*hai^h*	'crow'
tsoo	'great grand parent'	*tua'*	'son'	*tsoo"*	'beads'	*pui^h*	'eye'
nɨwɨ	'Indian'			*nɨwɨ"*	'liver'		
hai	'uncle'						

TABLE 3
Effect of final features in Shoshone (from Miller 1996)

Following morpheme	Series 1 (lenis)	Series 2 (fortis)	Series 3 (prenasal)	Series 4 (aspirate)
mother /pia/	nɨ=pia[2] [nɨ**ß**ia] 'my mother'		ɨ"=pia [ɨ**mb**ia] 'your mother'	
to have /pai/	tsoo-pai [tθoo**ß**ai] 'to have GG.parent'	tɨpa'-pai [tɨßa**pp**ai] 'to have pinenut'	tsoo"-pai [tθoo**mb**ai] ' to have beads'	hai^h-pai [hai**ɸ**ai] 'to have crow'
	nɨwɨ-pai [nɨwɨ**ß**ai] 'to have Indian'	tua'-pai [tua**pp**ai] 'to have son'	nɨwɨ"-pai [nɨwɨ**mb**ai] 'to have liver'	pui^h-pai [pui**ɸ**ai] 'to have eye'
	hai-pai [hai**ß**ai] 'to have uncle'			

[2] The symbol '=' indicates the boundary between a clitic and the morpheme on which the preceding clitic relies. The symbol '-' indicates a morpheme boundary.

In SNNP, the series 3 final feature corresponds to voiced fortis, and the Series 4 final feature has merged with the Series 2 feature[3]; thus, the language exhibits three final features—lenis, fortis, and voiced fortis feature, as mentioned earlier. The morphophonological alternations associated with these final features in SNNP are exemplified in Table 4. In these examples, the initial stop in the word /kuma/ 'husband' in its bare form is a plain voiceless stop /k/. In bi-morphemic words, however, the morpheme-initial stop alternates and is realized as lenis stop /g/, fortis stop /k/, or voiced fortis stop /gg/ when preceded by pronominal proclitic /i=/ (1st person, singular), /a=/ (3rd person, singular, indefinite), or /ɨ=/ (2nd person, singular), respectively. These examples illustrate how consonant alternation interacts with active morphology.

Table 4
SNNP pronominal proclitics with final features
(M = Miller 1996; N= Nichols 1974; T = Thornes 2003)

proclitic	Word	gloss	final feature
none	/kuma/	'husband'	
/i=/ (1 sg.) (N:13; T:77)	/i**g**uma/ (N: 13)	'my husband'	series 1 (lenis)
/aˈ=/ (3 indef.) (N:13)	/a**k**uma/ (N: 13)	'someone's husband'	series 2 (fortis)
/ɨˈˈ=/ (2 sg.) (M:694)	/ɨ**gg**uma/[4]	'your husband'	series 3 (voiced fortis)

The same contrasts occur root-medially as well (Table 5). That is, medial contrasts do not always reflect synchronic morphophonological alternation, as in the previous examples, but also exists as underlying contrasts.

TABLE 5
Example of root internal three-way contrast[5] in SNNP

word	gloss
/togaa/	'night'
/taka/	'arrowhead'
/kɨggɨ/	'leg'

The complete SNPP consonant phoneme inventory is given in Table 6. The three-way contrast is observed word medially for stops and affricates. Word initially these three series merge into a single fortis series. Nasals, by virtue of being always voiced, have a two-way

[3] An alternative analysis is that PN did not have series 4 feature, and no merger took place in NP (including SNNP). That PN did not have series 4 feature was suggested by Miller, Elzinga, and McLaughlin (2005) who demonstrated that the historical source of Central Numic the series 4 feature was an interaction of stress placement and the series 2 feature, which resulted in phonemic split of the original series 2 feature into series 2 and series 4 features in pre-Proto Central Numic.

[4] The second person singular proclitic carries series 3 final feature in Shoshone; therefore, in SNNP, it is expected to induce voiced fortis articulation at the following consonant.

[5] These words were obtained from the audio material LA114 as discussed in the Section 3.

contrast only. Interestingly, fricatives, which could have a three-way distinction just as well as other obstruents, lack the voiced fortis series. It is beyond the scope of the current study to investigate how this state of affairs has come about, but aerodynamic constraints disfavoring voiced fricatives (cf. Ohala 1983) might be the reason for the observed asymmetry in the consonant inventory.

TABLE 6
Consonant phoneme inventory[6] of SNNP[7]

	Bilabial	Alveolar	Palatal	Velar	Labialized Velar	Glottal
Stop:	p bb b	t dd d		k gg g	kw ggw gw	ʔ
Fricative:		s z				h
Affricate:		ts ddz dz				
Nasal:	mm m	nn n				
Glide:			y		w	

The fact that lenis and voiced fortis obstruents occur only word medially, and that there are many words that have a transparent morphological structure showing the morphophonemic status of their medial obstruents suggest that in words that are synchronically morphologically opaque, medial lenis and voiced fortis consonants might mark the presence of a previous morpheme boundary that has been lost, providing a clue for the historical development of these words. For example, /kɨggɨ/ is thought to be derived by reduplication (Miller 1967), and /pa-/ in /pagg^{w}i/ 'fish' is arguably attributable to /paa/ 'water'. From this, one might argue that many morphologically unanalyzable words are lexicalized remnants of previous multi-morphemic words created by cliticization, affixation, and compounding, and that today's root medial contrasts reflect once-active final features at morpheme boundaries. This hypothesis and its implication on the nature of SNNP fortis-lenis contrast will be considered in the section 6.3.

2.3. Issues on the consonant alternation

I have briefly reviewed the Numic consonant alternations in the light of their synchronic relationship with preceding morphemes and particular final features. I have also shown that any of the obstruents that surface as morphological alternatives in the context of final feature are also used as distinctive underlying phonemes in SNNP. Their phonological functionality and interaction with active morphology are well documented and understood. However, detailed phonetic descriptions of the medial consonants are still scarce. The only published instrumental

[6] Transcriptions used in this inventory is the one chosen by UC Berkeley Field Methods Class (Fall 2005-Spring 2006), which worked on the similar but distinct dialectal variation of NNP. I chose to use the same transcription so that the comparison can be made easily.

[7]There was one instance of long velar nasal [ŋŋ] (in word [joŋŋõõ], 'evening') in the speech material I used in this study. The medial consonant could be phonemic long velar nasal (/ŋŋ/) or nasalized allophonic variant of long voiced velar stop (/gg/). Due to the small set of data, I cannot determine phonemic status of this consonant. There were also few tokens I could not determine whether the medial consonant was voiced velar fricative [ɣ] or velar nasal [ŋ]. If these sounds represent phonemic /ŋŋ/ and /ŋ/, then these must be added to Table 6.

studies on NP speech sounds are by Waterman (1911) and Babel (2006), and more studies are needed.

Another issue surrounding SNNP obstruents is the use of controversial terms for their phonological labels. Ladeforged and Maddieson (1996) point out that the terms fortis and lenis are used in the linguistic literature with many different meanings, and agree with Catford's (1977) warning that "the terms tense/lax, strong/weak, fortis/lenis, and so on, should never be loosely and carelessly used without precise phonetic specification." This statement implies that there is considerable variation among consonants labeled as 'fortis' or 'lenis'. Thus, it is important that their descriptions include phonetic details so that is possible to compare them with 'fortis' and 'lenis' consonants in other languages.

With these issues in mind, I conducted a qualitative and quantitative study on the medial consonants. The focus of study is to investigate phonetic characteristics of each of the medial obstruent types; that is, to provide auditory impressions, acoustic properties, and physiological correlates to 'fortis', 'lenis', and 'voiced fortis' consonants. The details of the study, especially selection of acoustic parameters to investigate, were guided by previous work on NP speech sounds. Observations and findings presented in these studies will be discussed in the next section.

3. Previous studies on the Northern Paiute obstruents

Linguists working on NP have tried to specify the nature of its speech sounds as accurately as possible. A couple of instrumental studies have been conducted specifically to collect data on the physiological events and acoustic properties of the speech sounds. Other studies focus on other aspects of language—phonology, morphology and syntax—and have a limited treatment of speech sounds based only on auditory impressions. Descriptions from both types of studies offer valuable insights into the choice of temporal and spectral properties to investigate. In the next section I discuss some of the major characteristics of medial consonants that have been repeatedly pointed out in the literature.

3.1. Description of the Numic fortis/lenis contrast

Several studies provide descriptions of Oregon Northern Paiute (ONP) fortis/lenis contrast based upon auditory impressions of the sounds[8], and two recurring themes emerge from these studies. One is the use of multiple phonetic features in realizing the contrast. For example, Thornes notes that the articulatory targets for fortis and lenis are on idealized extremes along a continuum with respect to multiple articulatory variables and that "[a] fortis consonant is ideally an unvoiced geminate stop, whereas a lenis consonant is ideally a voiced fricative" (2003:29). Further, both Nichols (1974: 31) and Liljeblad (1996:24; 1950:130) added preaspiration and preglottalization as the optional or obligatory concomitant of fortis articulation. These descriptions suggest that the articulations referred to as 'fortis' and 'lenis' do not have straightforward phonetic correlates but are realized in a range of ways through the combination of several phonetic features.

[8] These studies must be taken with caution because ONP does not have voiced fortis series, and the reported characteristics, therefore, may or may not apply to the SNNP medial contrast. Due to a unique need to make three-way distinctions, SNNP fortis and lenis series might be articulated differently from the same series in the rest of the NP dialects. Still, these are varieties of the same language and the careful observations from the previous researchers would provide useful information in the investigation of SNNP speech sounds.

The other recurring theme is the lack of invariance of a physical correlate to a given series. Nichols (1974) pointed out that the fortis-lenis contrast is not realized by obligatory common difference but by relative difference of the feature value, as illustrated in the following example:

> "[I]n NP the systemic difference between lenis [m] and fortis [mm] may be represented in the speech of a single individual by either [m]:[mm] or [w̰]:[m], where an identical phonetic value [m] may represent either the lenis or the fortis series depending on the opposing sound" (31-32).

Thornes (2003:29) also pointed out the problem of analyzing NP sounds using binary features such as [+/- voice] because of the gradient nature of natural speech and the great deal of free variation.

Furthermore, previous reports differ from each other in terms of the relative importance of the features. For example, Liljeblad reports that the release of fortis stops may be voiced, though rarely, and maintains that voicing is not a distinctive feature in the northern dialects (1966:22). Nichols (1974:19) notes that gemination, preaspiration and preglottalization of fortis is optional and sporadic, but according to Liljeblad (1966:24) preaspiration becomes more common in the western region (though he does not specify the region). Waterman (1911: 33) reports that long consonantal duration for medial fortis stops is a stable feature, but according to Liljeblad, the key feature of fortis is not voicing or duration but forceful closure. Liljeblad describes the systematic difference between fortis and lenis as follows:

> "[I]n lenis transition the vowel and the following consonant gradually coalesce without any acute audible break between the two sounds. In fortis transition, on the other hand, the vowel, whether short or long, is abruptly cut by the quick and vigorous closure followed by voiceless and unaspirated release (in the southern dialects of NP also by a voiced unaspirated release)" (1966:23-24).

In summary, fortis and lenis series are typically characterized by a relative difference in one or more of the phonetic features, including the manner of articulation, voicing, and consonantal duration. Fortis stops are reported to have optional preaspiration or preglotatlization and give the auditory impression of a 'quick and vigorous' closure. In addition, there seems to be a considerable difference even within the same dialect groups in the way that speakers realize fortis-lenis contrasts.

3.2. Previous phonetic studies

There are, to my knowledge, only two published instrumental studies on NP speech sounds. The first instrumental study was conducted by Waterman (1911) with a male speaker of ONP. He obtained photographic images to examine lip positions for the vowel articulations and kymograph tracings of the oral airflow and glottal position during the production of words. From these physiological data, Waterman concluded that his speaker's medial stops are phonetically in two kinds—the fully voiced stops and voiceless stops that have double length of occlusion. Although Waterman uses the term "stop", the airflow data reveal that his speaker's voiced stops sometimes lack complete closure—some of the kymograph tracings show continuous airflow throughout the consonantal period. When the 'voiced' stops have complete closure, however, voicing ceases during the closure and it starts again from the point of stop release. Thus, his speaker's lenis stops are realized either as a fully voiced continuant or a voiceless unaspirated stop. Finally, in accord with Liljeblad's observation, Waterman's oral flow data confirms the presence of preaspiration[9] with fortis stops.

The other published phonetic study was done by Babel (2006) on the variety of NP spoken around Mono Lake, Bridgeport, Coleville, and Sweetwater. She investigated temporal property of the medial obstruents and found that closure duration[10] is significantly different among the three classes of obstruents: fortis has the longest closure duration; lenis, the shortest; and voiced fortis, intermediate duration but much closer to that of fortis. Her finding that the two classes of fortis exhibit longer consonantal durations than the lenis class agrees with the auditory impression described by the previous Numic scholars and Waterman's instrumental study.

4. Current study

This section discusses the methodological details of the acoustic study and presents findings from the qualitative and quantitative analysis on the three types of stops. From the review of previous scholarly works, it has emerged that fortis-lenis contrasts are phonetically realized by a combination of multiple features, including consonantal duration, relative timing of voice onset, manner of articulation, and optional preaspiration and preglottalization. Thus the specific goal of the quantitative anlysis is to examine the acoustic correlates of the auditory characteristics reported in the previous studies.

4.1. Data

The speech data used in the present study was the archived audio material housed in the Berkeley Language Center (BLC). The original material consists of analog recordings[11] of a female Paviotso speaker made by Margaret Wheat during her linguistic field work between the year 1950 and 1952, and subsequently deposited to BLC by Sidney Lamb. The material was then digitized by BLC at a 96 kHz sampling rate. According to Lamb's narration, the speaker was about 75 years old, born in Mill City, and lived in the Stillwater and Fallon area.

For the purpose of controlling the source of variability in consonantal articulation, from the 95 minutes of recording of words, phrases, and sentences, only nouns consisting of two or three syllables which were uttered in isolation were initially selected. The waveforms and spectrograms of each utterance were checked by eye, and the tokens for which acoustic signals were too faint for analysis were eliminated, leaving 179 tokens to be used for the subsequent analysis.

---

[9] Waterman describes the phenomenon as vowels having 'postaspiration'. Phonetically, what is described as preaspiration (of the postvocalic consonant) and postaspiration (of the preconsonantal vowel) are the same event: What these terms describe is a period of aspiration at the boundary of VC sequence.

[10] Since many tokens lack complete closure, she used, as a surrogate of closure duration, the duration between the offset of the preceding vowel, through medial consonant, to the onset of the following vowel.

[11] Detailed recording conditions such as the type of microphone, the location of the microphone, the type of recording device, conditions of the room, and etc. are not known.

TABLE 7
Prosodic structure type and frequency (N = 179)

Prosodic Structure		Frequency	Examples[12]	
2-syllable words (n = 146)	(C)VVCV	72	[paadu]	'daughter'
	(C)VCCV	36	[takka]	'arrowhead'
	(C)VCVʔ	19	[kuruʔ]	'stick'
	(C)VVCCV	8	[kaoppu]	'leg'
	(C)VCVV	7	[togaa]	'night'
	CVCCVV[13]	4	[kʷiʔnaa]	'eagle'
3-syllable words (n = 33)	(C)VCCVCV	14	[hakkʷabu]	'hail'
	(C)VCVCCV	6	[tɨbappi]	'pinenut'
	(C)VCVCV	5	[togabu]	'night'
	CVCCVVCVN[14]	3	[magguuhan]	'finger'
	(C)VVCVCV	3	[buusuna]	'grass'
	(C)VCCVCCV	1	[huggʷappu]	'wind'
	(C)VVCVCCV	1	[piidappu]	'fire'
	total	179		

4.2. Distributional properties

Prior to the acoustic analysis, some of the distributional properties of medial consonants were studied. All 179 tokens were transcribed phonetically, paying close attention both to the phonetic identity (i.e quality) and the length of the segments (i.e. quantity). The underlying representation for each word was determined by consulting published Numic studies, dictionaries, and cognate sets as a guide[15]. Next, prosodic structures for each token were recorded as, for example, CVVCV if the token was heard as having a sequence of onset, long vowel, short consonant, and short vowel, CVCCV if the token was heard as having a sequence of onset, short vowel, long consonant, and short vowel, and so on. The result is presented in Table 7. Among two-syllable words, most (108 out of 134 words) had a heavy-light (HL) syllable of either CVVCV or CVCCV. The next most frequent prosodic structure for two-syllable words was

[12] Numic consonants may be realized phonetically in various forms. For example medial lenis stop /b/ may be realized as [b~ß], and medial /d/ may be realized as [ɹ~ɾ~d]. In this paper, phonetic transcription of consonants should be regarded as broad transcription reflecting one common realization of the speech sound.

[13] Among 179 words, there were 4 occurrences of kwiʔnaa (with a variant form kwiʔnnaa), and this word is the only instance of CVCCVV form.

[14] Among 179 words, there were 3 occurrences of magguuhan, and this word is the only instance of CVCCVVCVN form. Although the NP does not have phonemic coda nasal, the presence of nasal segment in the speech is clearly recognized both from listening and acoustic signals in all three instances of magguuhan, and the transcription reflects this observation.

[15] Published material was primarily consulted in determining the underlying form of consonant and vowel quality. I transcribed vowel length as I heard it rather than as it was recorded in the published materials. This was because sometimes a given word is attested with different vowel length and I suspected that for some words vowel duration is subject to individual variation.

Determining a vowel's underlying length was somewhat problematic. Vowels in open syllable are generally longer than vowels in closed syllable. Accented vowels may sound longer than unaccented vowels. A word uttered in isolation might carry so-called 'list intonation' in which the final syllable may be accented. These are the potential source of errors in phonemic transcription.

CVCVʔ[16], with the light first syllable followed by a heavy second syllable (LH). Prosodic structure for 3-syllable words varied so much and the number of data was so small that it was not feasible to identify any 'preferred' patterns. Within the data set, the most frequent structure was a (C)VCCVCV form, with a light syllable followed by a heavy syllable and another light syllable (LHL).

4.3. Qualitative analysis

This section present some of the qualitative observations made from acoustic representations of the speech samples. Acoustic properties will be discussed in relation to auditory impressions of the sounds. When appropriate, attempts were made to infer articulatory events from acoustic data.

Figure 2 shows waveforms and spectrograms of tokens that have fortis, voiced fortis and lenis stops after a short vowel (panels 1-3) and a long vowel (panels 4-6). All figures show the speech signal of the identical temporal range (750 ms) and frequency range (0-5000 Hz). For the purpose of comparison, only 2-syllable words are used as examples. Under each figure, the SNNP word, English gloss, prosodic structure of the word, position of accented syllable (indicated by underline), and type of medial consonant are indicated.

Many of the phonetic characteristics discussed in the Sections 3.1 and 3.2 were confirmed in the acoustic data. First, SNNP stops have variable realizations, as indicated by the variable degree of stop bursts. Many of the voiced fortis stops have a very weak burst, as shown in panel 2. Although most of the fortis stops exhibit a clear burst, some of the fortis affricates lack a clear burst, as exemplified in panel 4, indicating that fortis affricates may be phonetically realized as fricatives. The auditory impression, however, is that these consonants are true affricates rather than fricatives. It seems that the rapid amplitude build-up at the vowel onset is a robust cue similar to the transient sound of a stop release even though there is no actual 'stop' involved in the articulation. In lenis articulation, the degree of constriction varies even more. Some tokens are realized as voiced stops, exhibiting a weak burst; some are realized as voiced fricatives; and others are realized even as approximants, lacking a clear boundary from the adjacent vowels.

Second, the timing of voice onset roughly correlates with the three-way medial contrast but also exhibit considerable variability. Fortis stops are uniformly voiceless and aspirated. Voiced fortis stops are, despite its name, realized as voiceless unaspirated stops: The period before the release is always voiceless, as shown in panel 2 and 5, and voicing begins at the time of or immediately after the release. While many of the voiced fortis stops have very short VOT (less than 10 ms) or zero VOT, a few tokens of voiced fortis stops exhibit considerably longer VOT. The waveform and spectrogram of one such case—an instance of /pagg^{w}i/ 'fish'—is shown in Figure 3. This medial voiced fortis stop has 24 ms of VOT (indicated by lines (a) and (b) in the figure), which is comparable to the shorter VOT exhibited by fortis stops. This proximity in VOT between voiced fortis and fortis make these two classes of sounds often hard

[16] Glottal stop is included in the representation of prosodic structure because all instances of this form have clear glottal stop ʔ word finally and the second syllable is heard as prominent as any other accented heavy syllables.

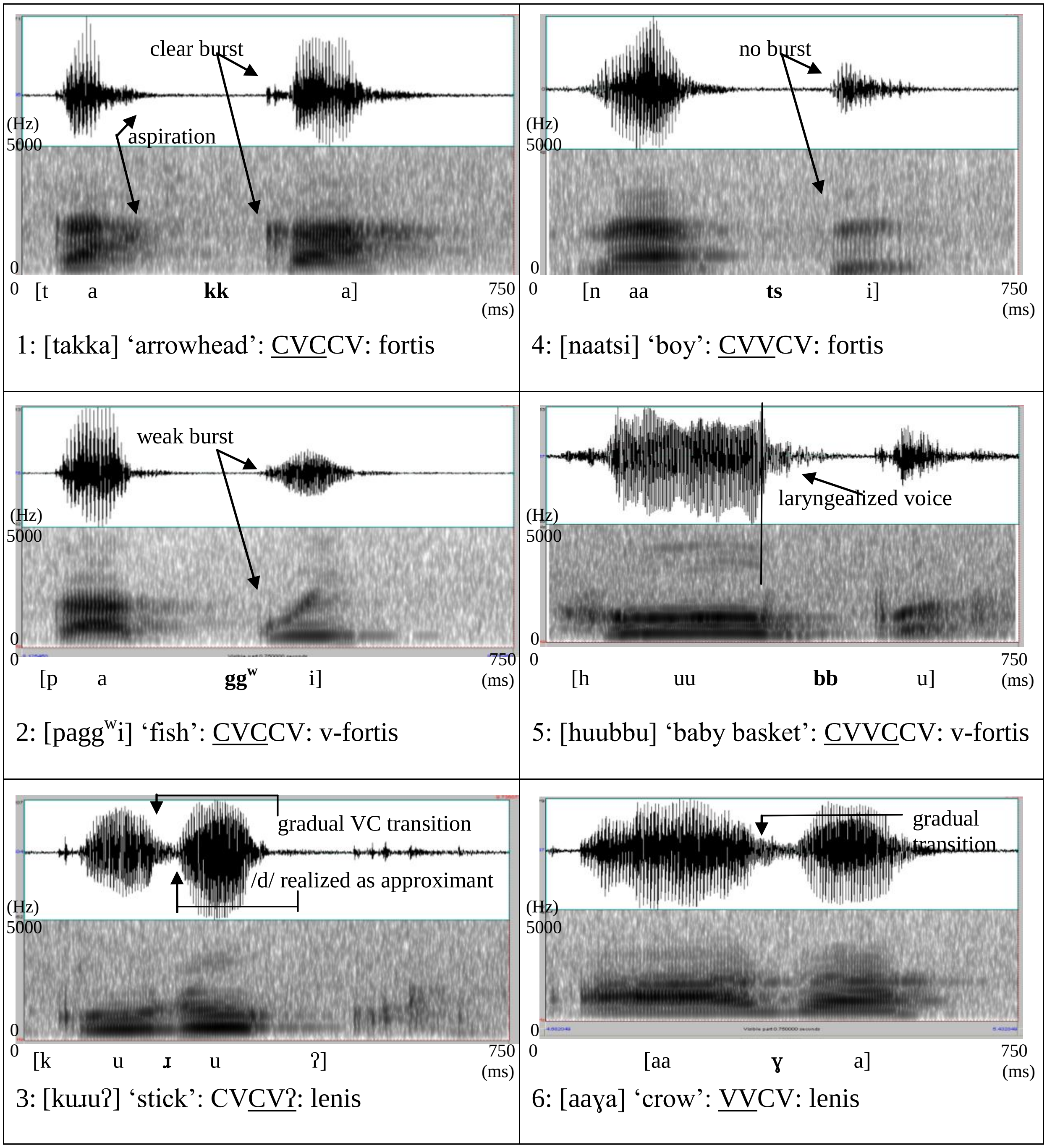


Fig. 2. waveforms and spectrograms of [takka] 'arrowhead', [pagg^{w}i] 'fish', [kuɹuʔ] 'stick', [naatsi] 'boy', [huubbu] 'baby basket' and [aaɣa] 'crow'.

to discriminate from each other. Fortis stops are not confusable, but voiced fortis stops are often confusable in its phonemic identity.

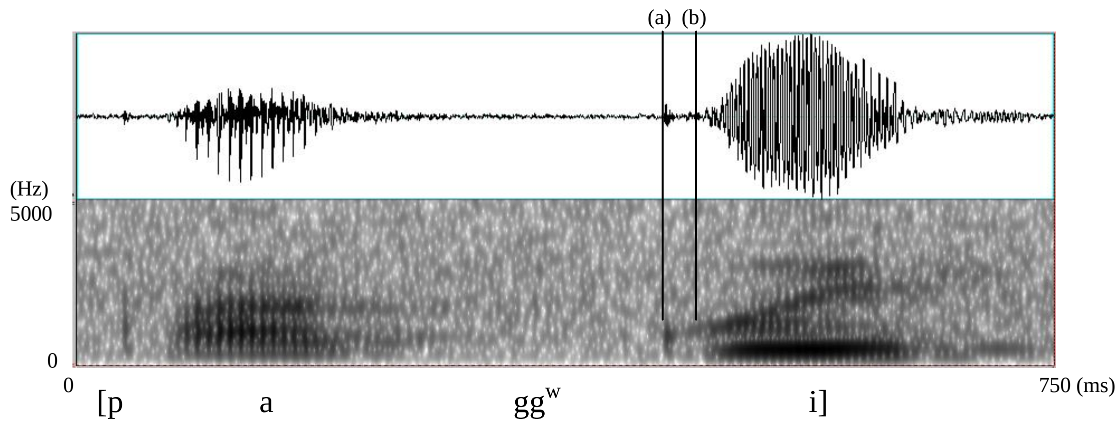


Fig. 3. Waveform and spectrogram of /pagg^{w}i/ ‘fish’

Third, the consonantal duration is much longer for the fortis and voiced fortis stops than for lenis stops. However, the acoustic evidence suggests that the long consonantal periods in the fortis and voiced fortis stops are not entirely occupied by complete closure. The spectrograms in panels 1, 2, 4, and 5 show weak vowel formants continuing for a considerable duration after the clearly audible portions of the vowels end. The corresponding waveforms show higher-than-baseline amplitude, indicating that the acoustic energy is present in the regions where weak vowel formants are observed. These signals suggest that there are temporal gaps between the offset of vocalic portion of the vowel and onset of oral closure for the following stop. This leads to the next point of observation.

Fourth, some of the fortis stops are pre-aspirated. An example is given in panel 1, which is repeated in Figure 4 for convenience. In Figure 4, the line (a) indicates the point at which the voicing of the vowel ends. During the period indicated between lines (a) and (b), the vocal tract is likely to be in the configuration of the vowel, but the glottis seems to start opening, producing weak formants with breathy voice. At the point indicated by line (b), the vowel formants are no longer visible, and only the aspiration noise continues.

The waveform reveals that there is a slight increase in the amplitude at (b). This is probably the same phenomena described in Waterman’s (1911) instrumental study. He observed, from the oral airflow trace, increased airflow at the end of vowels preceding the fortis stops. It can be inferred that during the production of the token in Figure 4, transglottal airflow increases at the end of the vowel, presumably due to the widened glottis.

Finally, a few tokens exhibited preglottalization before voiced fortis consonants (i.e. voiced fortis stops or nasals). One such case is shown in panel 5. In this figure, the vertical line indicates the point where the amplitude of the vowel abruptly decreases, and the following region shows irregularity, or cycle-to-cycle fluctuation, in amplitude and frequency—a typical acoustic manifestation of glottalized voice (Ohala, 1966; Titze, 1995; Gerratt & Kreiman, 2001; Gordon and Ladefoged, 2001; Hanson et al. 2001)—for about 80ms until the amplitude decreases to the baseline. Glottalized voice is associated with vocal folds that are tightly adducted but open enough along a portion of their length to allow for voicing (Gordon and Ladefoged, 2001) (cf. also Laver, 1980; Ní Chasaide and Gobl, 1995). SNNP glottalized voice toward the end of the vowel was probably produced with such tightly adducted vocal folds.

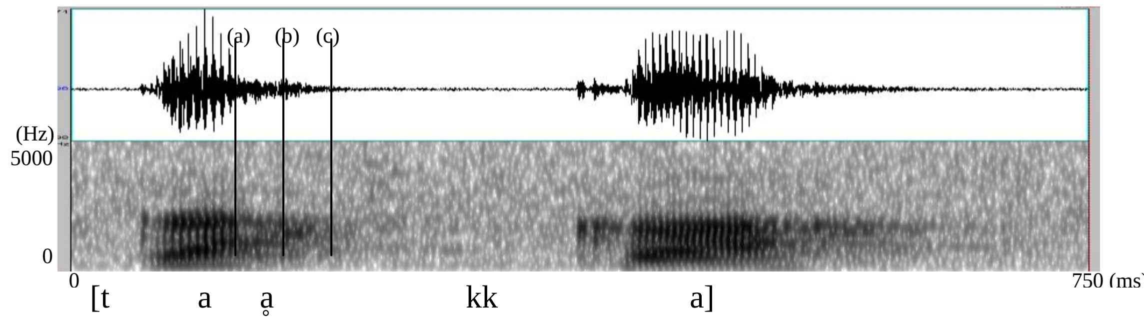


Fig. 4. Waveform and spectrogram of /taka/ ‘arrowhead’

In this section, some of the acoustic properties of SNNP stops were discussed. As often pointed out by Numic scholars, acoustic data confirms that there is a great deal of variability in realization of the stops, especially in the laryngeal features of voicing, aspiration, and glottalization. Therefore, the next step in investigation was to test the systematic use of acoustic properties given the range of variation. For this purpose, one-way ANOVA was performed on obtained measurements along each acoustic parameter. If the statistical test reveals a significant difference, then it will strongly support they hypothesis that speakers differentiate lenis, fortis, and voiced fortis articulation systematically by using that parameter.

4.4. Quantitative analysis

The focus in this section is statistical comparison of acoustic properties across medial consonant series[17]. In order to best control the source of variability of acoustic signals, the tokens that share the similar prosodic structure are selected. These are tokens consisting of two-syllables that have either of CVVCV, CVCCV, or CVCVʔ form – the three most frequently occurring prosodic structures among two-syllable words. Also, only tokens having medial stops and affricates (henceforth ‘stops’) are used for the analysis, resulting in 87 tokens to be used for the subsequent analysis.

4.4.1 Segmentation

Figure 5 and Figure 6 are examples of the segmentation. The words are /taka/ ‘arrowhead’ and /aaga/ ‘crow’, respectively. For all tokens, segmental boundaries for the first vowel (V1), aspiration following the vocalic portion of V1 (h1), a glottalized period following the period of modal voice in V1 (ʔ1)[18], medial consonant (C2), VOT of the second vowel (VOT 2), and the

[17] Since data is obtained from a single speaker, the statistical tests are used to obtain probability statements on the speaker’s linguistic behavior, not to make statements on the property of the language. The results may be generalized to the language so long as it is assumed that the speaker is representative of the entire speech community, as mentioned in the Introduction.

[18] The segment (h1) and (ʔ1) were marked only when there is clear evidence in the acoustic signal. Weakly implemented aspiration and glottalization thus might have been overlooked.

second vowel (V2) were identified. Since the beginning of the first consonant (C1) is not visible in the signal of a word uttered in isolation, C1 duration was not measured. For the vowel, its onset is set at the first glottal pulse where F1 is clearly visible, and its offset is set at the last glottal pulse where F1 is clearly visible[19]. Thus, the region where there is a weak vowel formant structure but no voicing (e.g. the regions corresponds to [ḁ] in Figure 5) is not included in the 'vowel'. Aspiration (h1) includes the region where there is a weak vowel formant and where only aspiration noise is present. This is seen in (h1) segment in Figure 5. Where there is no

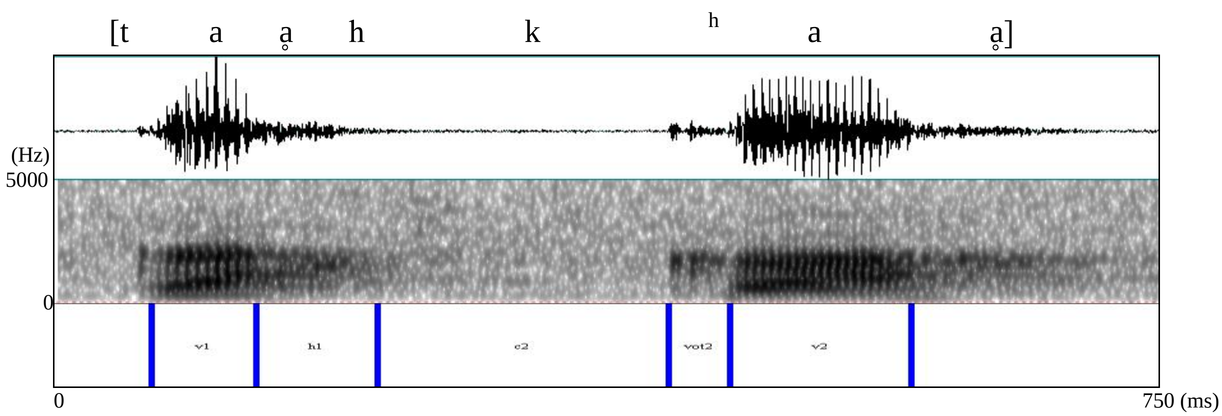


Fig. 5. Sample of segmentation (1): word is /taka/ 'arrowhead'

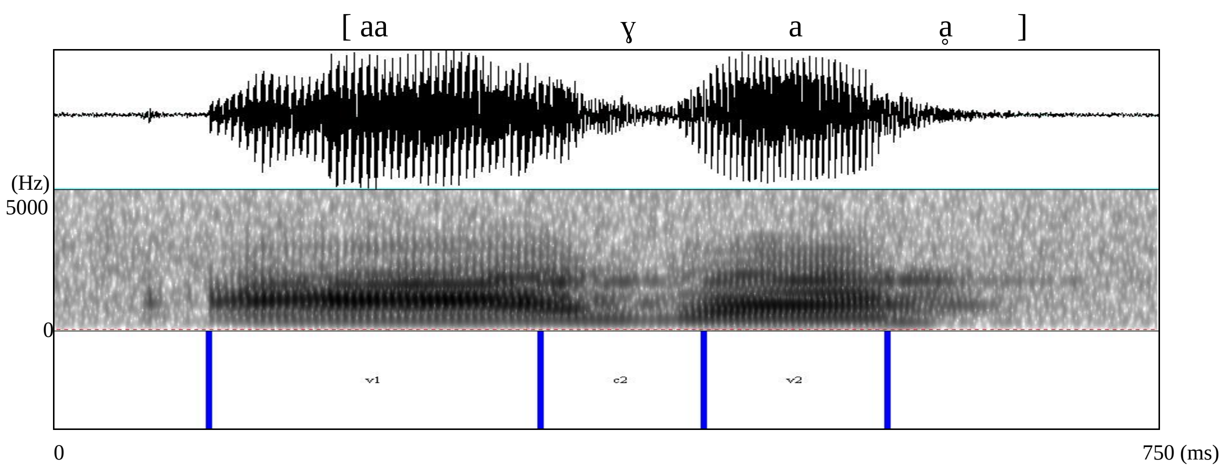


Fig. 6. Sample of segmentation (2): word is /aaga/ 'crow'

[19] This criterion was a modified version of the criterion used by Myers & Hansen (2005), where the vowel's boundary is set at the first/last glottal pulse where both F1 and F2 are clearly visible. In current study, this criterion is not practical because the audio generally lacks high frequency signal and F2 often becomes unclear in a middle of vocalic portion of the vowel (see, for example, V2 of paggwi 'fish' in Figure 2, Panel 2). Thus the criterion was modified slightly to better serve the purpose in this study.

clear boundary between vowel and consonant, as is often the case when the medial lenis is realized as a voiced fricative or approximant (e.g. /aaga/ 'crow' in Figure 6), the boundary was set in the middle of the F1 transition[20].

4.4.2 Manner of articulation and voicing

The first quantitative investigation was on the manner of articulation of the medial consonant and the timing of voice onset relative to the consonantal release. First, whether the medial stop is realized as a 'stop' or 'spirant' was recorded. The consonant was labeled as stop if a clearly visible stop burst was observed in the spectrogram; otherwise it was labeled as spirant. Stop tokens were further categorized as having positive VOT ('P-VOT', i.e. voicing starts after the stop release) or negative VOT ('N-VOT', i.e. voicing continues throughout the closure or begins before the stop release). Spirant tokens were categorized as fully voiced ('Voiced', i.e. voicing continues throughout the consonantal period) or voiceless. The frequency of occurrence of these four types of realizations in each of the three consonant series is provided in Table 8.

TABLE 8
Manner of articulation and Voicing (N=87).
(The values in parentheses indicate the percentage of occurrence within each consonant type.)

Medial Consonant	Realization of manner and voicing				Total
	Spirant		Stop		
	Voiced	Voiceless	N-VOT	P-VOT	
Lenis	40 (70%)	0	10 (19%)	7 (12%)	57
Fortis	0	6 (28%)	0	15 (71%)	21
Voiced fortis	0	2 (22%)	0	7 (78%)	9
Total	33	15	10	29	87

Among tokens with medial lenis stops, 40 out of 57 tokens (70%) exhibit spirantization, lacking a stop burst completely. This is the case where medial lenis is realized as either voiced fricative or approximant. 10 tokens (19%) exhibit a burst, and voicing continues before, at, and after the release. These tokens are voiced stops. Finally, 7 tokens (12%) exhibit a clear burst and zero or very small but positive VOT. These are voiceless unaspirated stops. Thus, although the majority of medial lenis is realized as either voiced fricative or stops, in a few cases medial lenis fails to maintain voicing to the end of the closure.

Among tokens with medial fortis stops, 15 tokens (71%) are realized with a clear burst and positive VOT, or as straightforward voiceless stops. There were six tokens (28%) that lack a burst, being realized as voiceless fricatives. These tokens demonstrate that SNNP fortis stops may be realized with incomplete oral closure.

Finally, as for tokens with voiced fortis stops, 7 tokens (78%) have a clear burst and short but positive VOT. Thus, the majority of voiced fortis are realized as voiceless unaspirated stops.

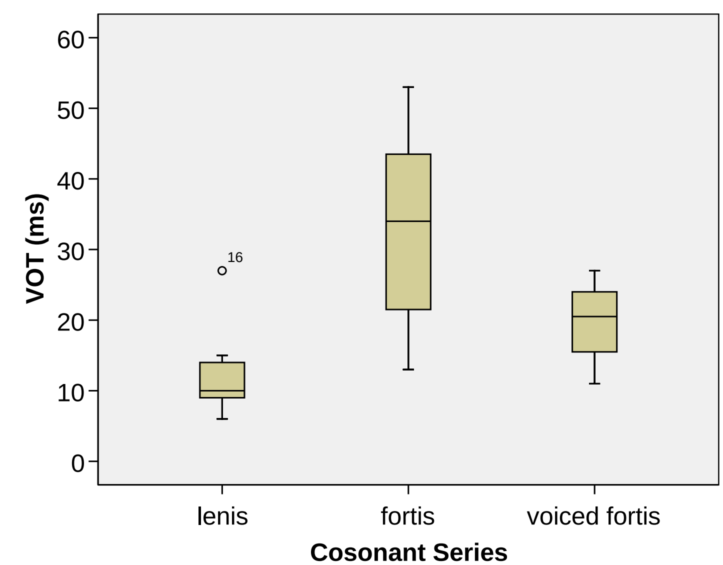
[20] This is the modified version of the criterion used by Klatt (1972:135 cited in Myers & Hansen 2005:322), where the vowel's boundary is set at halfway through the F2 transition.

Another 2 tokens (22%) lack a burst and voicing before vowel onset, being realized as voiceless fricatives.

In summary, in terms of voicing and manner of articulation, the majority of lenis stops are realized as voiced fricatives, while the majority of fortis and voiced fortis stops are realized as voiceless stops, with more aspiration for fortis stops and very short aspiration for voiced fortis stops. At the same time, all three series exhibited variation both in terms of manner of articulation and relative timing of voice onset.

4.4.3 VOT

As further investigation of voicing, VOT was measured in the 29 tokens which exhibit positive VOT. Figure 7 shows boxplots[21] of VOT of lenis, fortis, and voiced fortis stops. Mean VOT is greatest in the fortis series and smallest in the lenis series, as expected from the auditory impressions. Mean VOT of the voiced fortis series is intermediate and closer to that of the lenis than the fortis series. The boxplots show that lenis stops, if they have positive VOT, tend to have very short VOT, ranging from about 5 to 15 ms. VOT of fortis stops, on the other hand, have a wider distribution, ranging from about 15 to 55 ms. Despite the wide distribution of the VOT of fortis stops, VOT exhibits a categorical distribution between fortis and lenis stops: There is no overlap in VOT between the two series. VOT of the voiced fortis stops, on the other hand, overlaps VOT of both fortis and lenis stops. In particular, the range of VOT for the voiced fortis series is almost completely covered by the range of VOT of the fortis series.



Summary statistics

Medial consonant	n	Mean (ms)	Std. Dev (ms)
Lenis	7	13	7
Fortis	15	33	13
V- fortis	7	20	7
Total	29	26	14

Figure 7 Boxplots of VOT in fortis, lenis, and voiced fortis condition (N=29)
Summary statistics are provided in the right column.

[21] In the boxplots used in this paper:

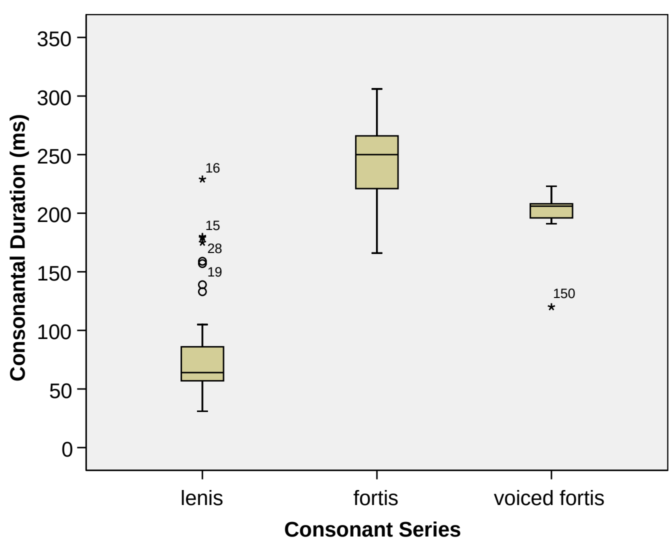
1) each box spans from the first quartile (i.e.the 25th percentile) to the third quartile (i.e. the 75th percentile);
2) a line in a box marks the median, which is indicated next to the box in parentheses;
3) lines extend from the box out to the smallest and largest observations, excluding outlier(s); and
4) an outlier is defined as a value more than 1.5 x IQR (Interquartile range: the 3rd quartile – the 1st quartile) above the 3rd quartile or below the first quartile.

The result of one-way ANOVA rejects the null hypothesis [F (2, 26) = 8.43, $p < 0.002$]. That is, at least one series of medial consonant has significantly different VOT from one other series. Scheffé's post-hoc[22] tests reveal that VOT of lenis and fortis are significantly different from each other and VOT of voiced fortis is not significantly different from either that of fortis or lenis. The ANOVA table and the post-hoc tests table are given in Appendix.

This result suggests that VOT varies according to stop series somewhat systematically, but there is a considerable overlap in VOT between the voiced fortis and both of the other series. This result confirms the difficulty of distinguishing medial consonant type based upon perceived voicing as reported in the previous studies.

4.4.4 Consonantal duration

The next measurement was on consonantal duration. This is a sum of durations of three parts—'h1/ʔ1', 'C2', and 'VOT 2'. This intervocalic period was thought to reflect perceived consonantal duration. This measurement was taken from all 87 tokens. Figure 8 shows boxplots of consonantal duration of lenis, fortis, and voiced fortis stops. Mean duration is greatest for fortis and smallest for lenis series. Mean duration of voiced fortis stops is in between but much closer to that of fortis stops. The difference of mean duration between these two series is 51 ms, which is much larger than mean VOT difference between these two series, thus the observed consonantal duration difference is not an artifact of VOT difference. The boxplots show that the lenis series exhibits a clearly different durational range from that of the fortis and voiced fortis series.



Summary statistics

Medial consonant	Mean (ms)	n	Std. Dev. (ms)
Lenis	76	55	38
Fortis	243	21	37
V- fortis	192	11	29
Total	131	87	82

Fig. 8. Boxplots of consonantal duration in fortis, lenis, and voiced fortis condition (N=87). Summary statistics are provided in the right column

[22] Scheffé's test is used for post hoc test throughout the paper.

The result of one-way ANOVA rejects the null hypothesis [F (2, 84) = 147.29, $p < 0.001$]. Subsequent post-hoc tests reveal that consonantal duration is significantly different across all three classes of medial consonants. The ANOVA table and the post-hoc tests table are given in Appendix.

These results agree with the findings of Babel (2006) both in magnitude of difference and direction. Her results show that the consonantal duration of the fortis series is almost three times longer than that of lenis series, and the duration of the voiced fortis series is in between but closer to that of the fortis series. Thus, consonantal duration seems to be a robust acoustic correlate to the medial consonant types.

4.4.5 Laryngeal involvement

Waveforms and spectrograms of each token were visually inspected to determine whether medial stops are accompanied by preaspiration or preglottalization[23]. Table 9 shows the number of tokens that exhibit either preaspiration or preglottalization at the vowel-consonant (VC) juncture. It shows that a majority of fortis stops are preceded by preaspiration, and the majority of voiced fortis stops are preceded by glottalization, while only a fraction of lenis tokens exhibit such laryngeal modulation. A statistical test was not performed because the number of observations was too small to perform a non-parametric test. However, the observed asymmetry in the distributional property between lenis series in one hand and fortis and voiced fortis series on the other hand suggests a potential association between the occurrence of laryngeal modulation and medial consonant types.

TABLE 9
Number of tokens exhibiting laryngeal feature at VC juncture (N=87)

Consonant type	Laryngeal feature			Total
	[h]	[ʔ]	none	
Lenis	0	7	50	57
Fortis	15	0	6	21
Voiced fortis	0	6	3	9
Total	15	13	59	87

4.4.6 Voice quality

As another way of examining the acoustic indication of preaspiration and preglottalization at the VC juncture, voice quality from the end portion of the vowel ('V1') was examined. The assumption is that when a vowel is followed by aspiration, the glottis starts to open before the end of the vowel, thus producing breathiness toward the end of the vowel. Similarly, it is

[23] There were many tokens where, due to very low sound-to-noise ratio, I was unable to determine the presence or absence, or the nature of laryngeal modulation from waveforms and spectrograms. I classified the tokens as "having laryngeal modulation" only when there was a clearly visible acoustic evidence. Thus, the result might be negatively biased; that is, the tokens that exhibit weakly implemented laryngeal modulation might be classified as not having such feature.

assumed that when the vowel is followed by glottalized voice the end portion of the vowel also exhibits weak glottalization.

Voice quality is often quantified from the vowel spectrum, by comparing the amplitude of the first harmonic to that of a higher harmonic such as the second harmonic (H1-H2) or the first-formant peak (H1-A1) (Ladefoged et al. 1987; Ní Chasaide & Gobl 1997; Hanson & Chuang 1999; Gordon & Ladefoged 2001; Hanson et al. 2001). These acoustic parameters are shown in Figure 9. Breathy voice involves a long open phase and slow closing or incomplete closure of the vocal folds between each successive glottal opening (Ohala, 1966; Laver 1980; Ní Chasaide & Gobl 1997), which boosts the lowest harmonic and attenuates the harmonics at mid and high frequencies and results in a sharp spectral tilt in the voice spectrum (Titze, 1994; Ní Chasaide & Gobl 1997; Hanson & Chuang 1999; Gordon & Ladefogrd 2001; Hanson et al.2001). On the other hand, glottalized voice involves a short open phase, rapid closure, and long closed period in each glottal cycle (Ohala 1966; Ní Chasaide & Gobl 1997). This reduces the levels of lower harmonic relative to the higher harmonics (Ní Chasaide & Gobl 1997; Gordon & Ladefoged, 2001), resulting in much flatter spectral tilt on voice spectra. In vowel spectra H2 and A1 are higher than H1 due to amplification by F1, yet the relationship that H2 and A1 relative to H1 are lower for breathy voice than for glottalized voice would remain.

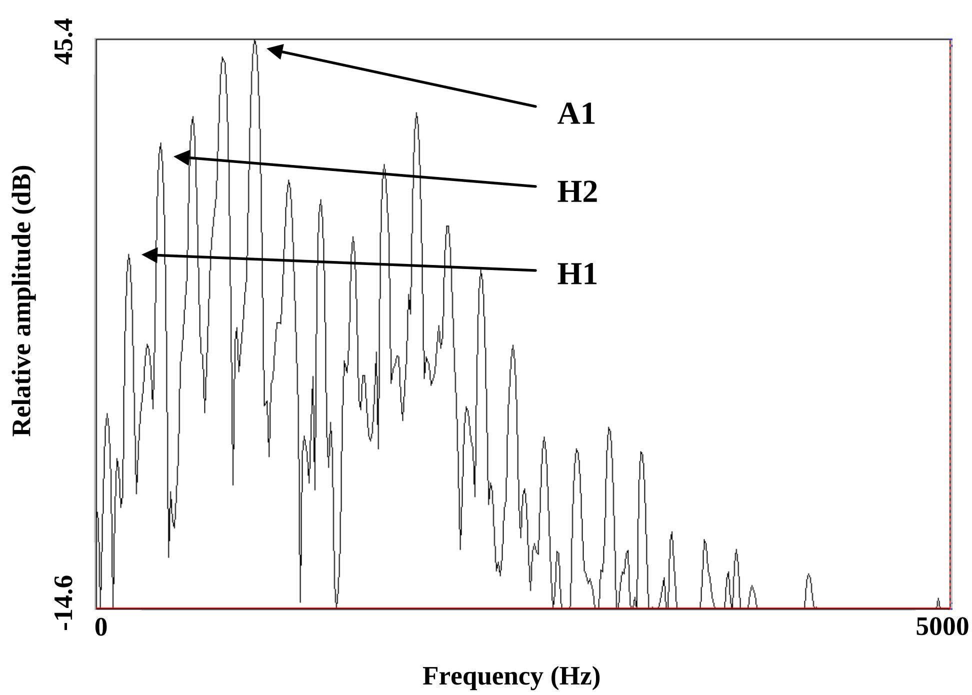


Fig. 9. An example of vowel spectrum, showing the amplitude of the first harmonic (H1), second harmonic (H2) and the first formant peak (A1).

The constraint of this method is that the vowel has to be the same across tokens. F1 differs across vowels; and so does the degree of influence of F1 on the H2. For example, F1 is inversely correlated with vowel height, thus closer to the second harmonic for high vowels than for low vowels given the same voice quality and F0. As a result, H2 would be higher for high vowels because of the stronger boost by F1 in high vowels than in low vowels. Therefore, in this study only the tokens that have the low vowel [a] preceding the medial consonant—the tokens that would exhibit smallest degree of F1 influence on H2—were selected.

There were 25 such tokens--13 tokens with lenis stops; 7, with fortis, and 5 with voiced fortis stops--selected for the measurement. For these tokens, H1-H2, and H1-A1 values were taken from the last 30 ms period of the 'V1' segment. The reason why this measurement was not taken from 'h1' or 'ʔ1' segment but rather taken from the end of the 'V1' segment was that obtaining vowel spectra with clearly identifiable H1 and H2 requires the signal to exhibit an identifiable glottal pulse. This condition excludes the 'ʔ1' segment and noise portion of the 'h1' segment: In the former, due to irregularity of the signal each glottal pulse may not be identified; and in the latter, a glottal pulse is not present. This is a stringent method because measurement is made where a large effect of aspiration or glottalization is not expected.

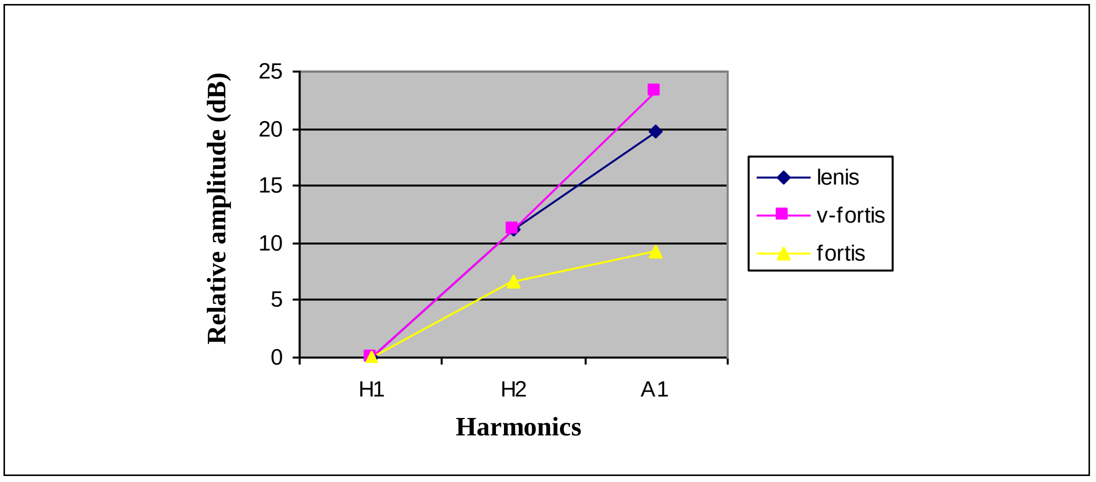
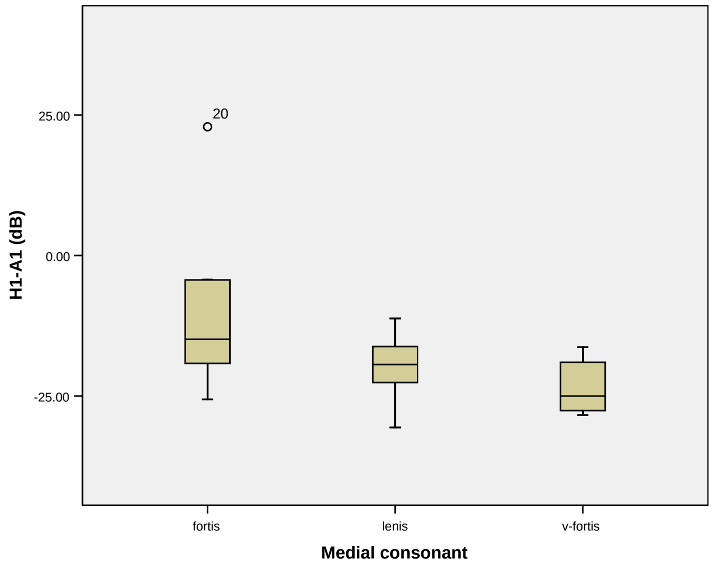


Fig. 10. Relative amplitude at first harmonic ('H1'), second harmonic ('H2') and F1 peak ('A1') compared to H1, obtained from vowels in lenis, voiced fortis, and fortis stop environments

Figure 10 shows line plots of mean relative H1 (H1-H1), relative H2 (H2-H1), and relative A1 (A1-H1) values, showing the amplitude increase at H2 and A1, in three different consonantal environments. It shows that, comparing with vowels before lenis stops, vowels before fortis stops exhibit less boost at both H2 and A1. This indicates that the vowel is produced with the voice quality that involves relatively lower acoustic energy in higher frequency components. Therefore, I interpret this to mean that vowels before fortis stops generally have a more breathy quality than vowels before lenis stops.

Vowels before voiced fortis stops, on the other hand, exhibit more boost at A1 compared with vowels before lenis stops, indicating that these vowels are produced with the voice quality that has relatively greater acoustic energy in higher frequency components. I interpret this to mean that vowels before voiced fortis stops have a more glottalized quality than vowels before lenis stops. However, the fact that H1-H2 is comparable in lenis and voiced fortis conditions, and the rather small difference of H1-A1 values between these two conditions, suggest that glottalization is either weakly or inconsistently implemented.



Summary statistics

Medial cons	Mean (dB)	N	Std. Dev. (dB)
fortis	-9.24	7	16.24
lenis	-19.75	13	5.26
v_fortis	-23.26	5	5.36
Total	-17.51	25	10.68

Fig. 11. Boxplots of H1-A1 in fortis, lenis, and voiced fortis condition (N=25)
Summary statistics are provided in the right column

The results from statistical tests support the above interpretation. For H1-H2, there is no significant consonant type effect [F (2, 22) = 1.350, p = 0.280]. Although we see from Figure 10 that H1-H2 values are higher in fortis conditions than in voiced fortis and lenis conditions, given the sensitivity of the test, the observed difference does not reach statistical significance. For H1-A1, on the other hand, there was a significant consonant type effect [F (2, 22) = 3.841, p < 0.05]. Subsequent tests reveal that mean H1-A1 in fortis conditions is significantly different from that in voiced fortis and lenis conditions (see Figure 11). The ANOVA table and the posthoc tests table are given in the Appendix. These results reinforce the interpretation that preaspiration systematically accompanies fortis stops. Preglottalization accompanies voiced fortis stops, but does so less regularly or with a smaller degree of glottalization. However, it should be noted that the method employed in this study is rather a stringent one because the evidence of pre-aspiration and preglottalization is taken from the segment 'preceding' the aspiration and glottalized portion of voice. Therefore, even a small effect would be a strong evidence for the presence of preaspiration and preglottalization.

5. Summary of the result

Qualitative and quantitative analysis on the acoustic data confirmed many of the phonetic characteristics described in the previous works on NP. As for the manner of articulation, the majority of lenis segments are realized as fricative or approximants, and the majority of fortis and voiced fortis are realized as stops. VOT is significantly different between lenis and fortis, but VOT of voiced fortis overlaps considerably with that of lenis and does so completely with that of fortis. The difference between fortis and voiced fortis and between voiced fortis and lenis in terms of VOT is, therefore, subtle. Voiced fortis is not realized phonetically as a voiced stop but as a voiceless unaspirated stop, in which voicing begins only after the stop release.

Consonantal duration is one of the most robust acoustic characteristics differentiating three classes of stops. The difference is particularly large between lenis stops on the one hand and fortis and voiced fortis stops on the other. In terms of the VC transition and voice quality, lenis stops are characterized by having a smooth VC transition with constant voice quality through the transition. Fortis stops tend to have aspiration at the VC juncture, which seems to contribute to an auditory impression of a sharp boundary between vowel and medial consonant. A few voiced fortis stops have glottalization at the VC juncture, also creating the impression of clear boundary between vowel and consonant. Since this is a study of single speaker's data, I would like to present the results as tentative. Further study with more data will be required to support them.

6. Discussion

In this last section, some of the characteristics of the three-way contrast will be discussed. Subtle distinctions between fortis and voiced fortis stops will be related to the merger of the two series in other dialects. Also, that the contrasts are not realized by a release feature of stops but by entire vowel-consonant sequence will be discussed in the context of the typology of stop contrasts. Finally, a hypothesis on the historical origin of the fortis and voiced fortis stops will be proposed.

6.1. SNNP three-way contrast

Results from qualitative observations and acoustic measurements have demonstrated that SNNP medial contrast is a unique three-way contrast which correlates with several phonetic properties, including a mode of vowel-termination, consonantal duration, manner of articulation, and VOT. Involvement of multiple phonetic features in the realization of phonological contrast is in accordance with previous reports by Nichols (1974) and Thornes (2003) on phonetic realizations of medial consonants in NP.

Acoustic measurements have shown that there is a good deal of overlap between the fortis series and voiced fortis series in their acoustic properties. In particular, very small mean VOT difference between fortis and voiced fortis stops and the fact that observed VOT values of voiced fortis stops is completely within the range of VOT of fortis stops indicate that these two series of consonant are very similar in their voicing property.

On the other hand, the contrast between the lenis series and the fortis series is very clear. Mean consonantal duration for fortis stops is more than three times longer than that of lenis stops. Moreover, the former are accompanied by a clear vowel-consonant boundary while the latter have no clearly recognizable boundary. This suggests that Liljeblad's (1966) description of the transitions between the preceding vowel and the consonant as being very different between fortis stops and lenis stops applies to SNNP medial contrast.

Very small VOT difference between the fortis series and the voiced fortis series was partly due to the failure to maintain voicing during consonantal closure in the voiced fortis series. In all tokens that have medial voiced fortis stops voicing is terminated before the stop release. The fact that voicing begins immediately after stop release suggests that the vocal folds are probably in the position of voicing (i.e. moderately approximated) (cf. Vencov, 1968) just before the release. Thus the cessation of voicing toward the end of the closure is likely to be caused by adverse aerodynamic conditions, or heightened intraoral pressure due to the long stop closure and subsequent decrease of the transglottal pressure drop.

The language could utilize some mechanism of maintaining voicing such as prenasalization, shortening of the closure duration, etc. (Ohala, 1983) in the case of adverse aerodynamic conditions; however, SNNP opted not to resort such measures and maintained the contrast between the fortis and voiced fortis series with rather subtle differences of VOT. This could be one of the underlying conditions of merger of voiced fortis and fortis to fortis that had taken place in other dialects of NP. The three-way contrast would have been more robust if voiced fortis were fully voiced, or fortis had distinctively long VOT.

6.2. SNNP fortis-lenis contrast

In SNNP, one of the most robust physical correlates of the fortis-lenis contrast is discontinuity in the acoustic signal at the vowel-consonant boundary in the fortis and voiced fortis series and lack of such discontinuity for the lenis series. This discontinuity is primarily achieved by long consonantal duration and further reinforced by optional co-occurrence of preaspiration and prelaryngealization in the fortis and voiced fortis articulations. Although these are opposite configurations—one involves glottal abduction and the other involves tight adduction—crucially, they achieve the same end of terminating vocal fold vibration and thus reinforcing a clear discontinuity in the acoustic signal. In this respect, one might argue that preaspiration and preglottalization are utilized as means rather than ends, and medial contrasts are realized on an entire vowel-consonant sequence rather than on a single segment.

Treating the SNNP medial contrast as a property of a vowel-consonant sequence also explains why the three-way contrast occurs only word-medially but not initially: The realization of the contrast crucially relies on the presence of the preceding vowel. An interesting parallel is observed in languages such as Icelandic and Faroese that use preaspirated stops as phonologically distinctive series against plain stops (Maddieson and Ladefoged, 1996; Helgason, 2002). Their stop series contrast occur only word medially and finally, but not initially. What is common in these languages is that distinctive laryngeal contrast occurs at the vowel-consonant boundary rather than consonantal release.

The idea that some languages use the vowel-consonant boundary rather than consonantal release as a target where stop contrast is realized is discussed by Jansen (2004). Citing Steriade 1997 and Helgason 1999, Jansen notes that preaspirated fortis stops in Icelandic, Faroese, Norwegian, Swedish, and English suggest that "Voice Termination Time" (VTT, i.e. the relative timing of the onset of an obstruent and the offset of voicing[24]) can be used as a cue to the fortis. Thus among languages that have a two-way contrast in the stop series there may be a variation in the temporal location where the phonetic feature associated with the contrast is realized. Languages may target either the stop release or vowel-consonant juncture for this purpose, and SNNP obviously targets the latter.

6.3. On the origin of voiced fortis

The fact that some of the fortis stops exhibit preaspiration and some of the voiced fortis stops exhibit preglottalization leads to the speculation that the fortis and voiced fortis series might have

[24] This corresponds to 'Consonantal Duration' minus 'VOT2' in the current study.

been derived, historically, from VhC[25] and VʔC sequences, respectively. As for the voiced fortis series, this speculation is strongly motivated by the fact that the fortis sonorant, or long sonorant, often varies freely with a glottal stop-consonant sequence. For example, [kwiʔna] and [kwinna] 'eagle' are in free variation, and so are [moʔmogoni] and [mommogoni] 'woman (pl.)'. It could be the case that the segment-like realization of glottal stop has survived only between vocalic segments where it can be clearly heard, and before stop where the presence of glottal stop is hard to be detected, it has been reduced to become a weak feature.

Also, the above speculation matches the morphology of the language. There are many words that have been successfully decomposed into component monosyllabic morphemes (see for example, Natches 1923; Miller 1967; Nichols 1974; Poldervaat 1989; Stubbs 2006). Poldervaart (1989) maintains that Northern Paiute words are basically formed by concatenation of monosyllabic morphemes. Assuming that Poldervaart's claim is true, today's root-medial contrast can be viewed as a reflection of the historical development of NP words.

Further, both VhC sequence and VʔC sequence naturally follow from various forms that NP rhymes can take. From the results of instrumental study, Waterman (1911) concludes that NP final syllables are closed either by aspiration or by a glottal, yielding Vh, and Vʔ as possible final forms. Acoustic data obtained in the present study confirms these two variations in addition to V(V) as the possible syllable-final forms.

When a morpheme with a final form V(V), Vh, and Vʔ is followed by another morpheme, the derived form will have three different types of medial sequence: V(V)C, VhC, and VʔC. That fortis and voiced fortis are sometimes accompanied by preaspiration and preglottalization, that Mono has hC sequence corresponding to MLNP fortis, that long sonorants often freely alternate with ʔC sequence, and the possibility that the multi-syllabic words are the outcome of historical morpheme concatenation seem to provide enough support to turn this speculation into a reasonable hypothesis: fortis and voiced fortis have been derived, historically, from VhC and VʔC sequences, respectively.

This hypothesis has several explanatory merits. One is the ability to explain the observed difference in consonantal duration between the fortis and voiced fortis series on one hand and lenis series on the other. In every case of morpheme concatenation involving final h or ʔ in the preceding morpheme, the first syllable of the derived word would be a heavy syllable. Although internal morphology is no longer transparent in today's forms and thus h and ʔ are reduced to be non-normative features, prosody has been kept. In this scenario, the long consonantal duration in fortis and voiced fortis can be seen as a remnant of the original heavy syllable.

Another merit is the ability to explain the observed timing of the voicing in the three stop series: Lenis stops have voicing all the way through the consonantal period, fortis stops have

[25] Helgason (2002) demonstrates for the sound change in the opposite direction, where non-normative preaspiration has become normative feature in Nordic languages including standard Swedish, Faroese, and Icelandic. Given two closely related languages exhibiting phonemic hC and subphonemic preaspiration, either of the two features (i.e. normative or non-normative feature) can be a source of the other. In the case of MLNP, possibility that VʔC sequence is the source of voiced fortis VC sequence encourages the hypothesis that VhC is the source of fortis VC sequence.

relatively short lag voicing after consonantal release and voiced fortis stops have voicing immediately after stop release. Failure to maintain voicing in the voiced fortis stops can be explained by the relative difficulty in initiating vocal fold vibration compared to maintaining it (Lindqvist, 1972). To initiate voicing, the vocal folds need to be approximated and there must be sufficient transglottal pressure drop, estimated at approximately 2 to 3 cm H2O (Lindqvist, 1972; Ohala and Riordan, 1979). During the stop closure, intraoral pressure builds up behind the closure and transglottal pressure drop decreases; thus if the vocal folds are not already set in vibration, as in the case of word initial stops, voicing is less likely to begin prior to stop release. In the case of voiced fortis stops, where vocal fold vibration is deliberately terminated by glottalization, it is also unlikely that voicing resumes prior to stop release. In the case of lenis stops, continuous air leakage through incomplete oral closure facilitates maintenance of necessary transglottal pressure drop, and thus voicing continues all the way to the following vowel.

The explanation of the relatively short VOT in fortis stops rests on the relative timing of glottal gesture and oral gesture. Kingston (1990: 427) reports that during the sequence of a vowel and a cluster of a voiceless and a voiceless stop, peak glottal opening does not occur at the center of a fricative or stop but occurs in a middle of the cluster. Therefore, the glottal opening at the time of stop release is narrower in fricative-stop clusters such as /sp-/ and /st-/ than in the stop alone, leading to shorter VOT in the former than latter. The same articulatory timing may be employed in the SNNP fortis stops, which is preaspirated; that is, peak glottal opening occurs relatively earlier to oral articulation, contributing to relatively short lag in voicing after stop release.

Assuming that today's fortis and voiced fortis series were developed from hC and ʔC sequences, the observed time course of voicing can be explained as a natural consequence of the aerodynamic and articulatory constraints on voicing. This hypothesized historical make up of the medial contrast thus has great advantages in explaining the patterning of the SNNP sounds.

References

Babel, Molly. 2006. The lenis and fortis contrast in Mono Lake Paiute: A socio-historic phonetic study. paper presented at the *Friends of Uto-Aztecan Conference*. Salt Lake City, Utah.

Catford, J.C. 1977. *Fundamental Problems in Phonetics*. Bloomington: Inidiana UP.

Gerratt, Bruce R. and Kreiman, Jody. 2001. Toward a taxonomy of nonmodal phonation. *Journal of Phonetics*. 29: 365-381.

Goddard, Ives, ed. 1996. *Handbook of Northern American Indians*, vol. 17, *Languages.* Washington: Smithsonian Institution.

Gordon, Matthew and Ladefoged, Peter. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics*. 29: 383-406.

Hanson, Helen M. and Erika S. Chuang. 1999. Glottal characteristics pf male speakers: Acoustic correlates and comparison with female data. *Journal of Acoustical Society of America.* 106 (2): 1064-1077.

Hanson, Helen M., Kenneth N. Stevens, Hong-Kwang Jeff Kuo, Marilyn Y. Chen, and Janet Slifka. 2001. Towards models of phonation. *Journal of Phonetics.* 29: 451-480.

Helgason, Petur. 1999. Phonetic preconditions for the development of normative preaspiration. *Proceedings of the 13th International Congress of the Phonetic Sciences.*

Helgason, Petur. 2002. *Preaspiration in the Nordic Languages: Synchronic and diachronic aspects*. Doctoral Dissertation. Stockholm University.

Jansen, Wouter. 2004. *Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology Approach to English*, *Hungarian, and Dutch*. Doctoral Dissertation, University of Groningen.

Kingston, John. 1990. Articulatory binding. In J. Kingston & M. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, pp. 406-434. Cambridge: Cambridge University Press.

Ladeforged, Peter, Maddieson, Ian, and Jackson, Michael. 1987. Investigating phonation types in different languages. In O. Fujimura (ed.), *Vocal physiology: voice production, mechanisms and functions,* pp. 297-317. NewYork: Raven.

Ladefoged, Peter and Ian Maddieson. 1996. *Sounds of the world's languages*. Oxford: Blackwells.

Langacker, Ronald W. 1976. A Note on Uto-Aztecan Consonant Gradation. *International Journal of American Linguistics*. 42: 374-379.

Laver John. 1980. The Phonetic Description of Voice Quality. Cambridge, MA: Cambridge University Press.

Liljeblad, Sven. 1950. Bannack I: Phonemes. *International Journal of American Linguistics*. 16: 126-131.

Liljeblad, Sven. 1966. Northern Paiute Manual I: Grammatical Sketch of the Northern Dialects. Pocatelo: Dept. of Anthropology, Idaho State University.

Maddieson, Ian and Ladefoged, Peter. 1985. "Tense" and "lax" in four minority languages of China. *Journal of Phonetics*. 13: 433-454.

Miller, Wick. 1967. Uto-Aztecan Cognate Sets. *University of California Publications in Linguistics,* vol. 48. Berkeley: University of California Press.

Miller, Wick. 1996. Sketch of Shoshone, a Uto-Aztecan Language. *Handbook of North American Indians*, vol. 17, *Languages*, ed. by Ives Goddard, pp. 693-720. Washington D.C.: Smithsonian Institution.

Miller, Wick, Elzinga, Dirk, and McLaughlin, John E. 2005. Preaspiration and germination in Central Numic. *International Journal of American Linguistics*. 71: 413-444.

Myers, Scott and Hansen, Benjamin B. 2005. The origin of vowel-length neutralization in vocoid sequences: evidence from Finnish speakers. *Phonology*. 22: 317-344.

Natches, Gilbert. 1923. Northern Paiute Verbs. *University of California Publications in American Archaeology and Ethnology*. 20: 245-259.

Ní Chasaide, Ailbhe and Christer Gobl. 1997. Voice Source Variation. In Hardcastle W, Laver J, eds. *The Handbook of Phonetic Sciences*. MA.: B. Blackwell: 427-461.

Nichols, Michael Porter. 1974. *Northern Paiute historical grammar*. Doctoral Dissertation, University of California.

Norris, Evan J. 1986. *A Grammar Sketch and Comparative Study of Eastern Mono*. Doctoral Dissertation. University of California, San Diego

Ohala, John J. 1966. A New Photo-Electric Glottograph. *UCLA Working Papers in Phonetics* 4: 40-52.

Ohala, John J. and Riordan, Carol J. 1979. Passive vocal tract enlargement during voiced stops. In: Wolf, J.J. & Klatt, D.H. (eds.), *Speech communication papers*. (pp. 89-92). New York: Acoustical Society of America.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In MacNeilage P.F. (ed.), *The production of speech* (pp. 189-216). New York: Springer-Verlag.

Poldervaart, Arie. 1989. Northern Paiute Subanalysis. paper presented at the Friends of Uto-Aztecan Conference. Tucson, Arizona.

Sapir, Edward. 1930. The Southern Paiute language. *Proceedings of the American Academy of Arts and Sciences* 65: 1-296.

Steriade, Donca. 1997. Phonetics in phonology: the case of laryngeal neutralization. Ms., UCLA.

Stubbs, Brian D. 2000. More Palatable Reconstructions for Uto-Aztecan Palatals. *International Journal of American Linguistics*. 66: 125-137.

Thornes, Timothy Jon. 2003. *A Northern Paiute grammar: with texts*. Doctoral Dissertation, University of Oregon.

Titze, Ingo R. 1995. Definitions and nomenclature related to voice quality. In Fujimura O. and Hirano M (eds.), *Vocal fold physiology: voice quality control* (pp. 335-342). San Diego: Singular.

Vencov, A.V. 1968. A mechanism for production of voiced and voiceless intervocalic consonants. Zeitschrift fur Phonetik, Sprachwissenschaft und Kommunikationsforschung. 21: 140-144.

Waterman, Thomas Talbot. 1911. The phonetic elements of the Northern Paiute language. *University of California Publications in American Archaeology and Ethnology*. 10: 13-44.

Appenidix
Table A: ANOVA table
Dependent variable: VOT
Independent variable: medial consonant type (3 levels: lenis, fortis, voiced fortis)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.002	2	.001	8.430	.002
Within Groups	.003	23	.000		
Total	.005	25			

Table B: Scheffé's Multiple Comparisons Table
Dependent variable: VOT
Independent variable: medial consonant type (3 levels: lenis, fortis, voiced fortis)

(I) consonant	(J) consonant	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
lenis	fortis	-.02035(*)	.00515	.003	-.0338	-.0069
	v_fortis	-.00704	.00705	.614	-.0255	.0114
fortis	lenis	.02035(*)	.00515	.003	.0069	.0338
	v_fortis	.01332	.00633	.132	-.0033	.0299
v_fortis	lenis	.00704	.00705	.614	-.0114	.0255
	fortis	-.01332	.00633	.132	-.0299	.0033

* The mean difference is significant at the .05 level.

Table C: Scheffé's Table for homogeneous subset
Dependent variable: VOT
Independent variable: medial consonant type (3 levels: lenis, fortis, voiced fortis)

cons	N	Subset for alpha = .05	
		1	2
lenis	7	.0127	
v_fortis	4	.0198	.0198
fortis	15		.0331
Sig.		.538	.124

Means for groups in homogeneous subsets are displayed.
a Uses Harmonic Mean Sample Size = 6.528.
b The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Table D: ANOVA table
Dependent variable: Consonantal Duration
Independent variable: medial consonant type (3 levels: lenis, fortis, voiced fortis)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.455	2	.228	147.293	.000
Within Groups	.130	84	.002		
Total	.585	86			

Table E: Scheffé's Multiple Comparisons Table
Dependent variable: Consonantal Duration
Independent variable: medial consonant type (3 levels: lenis, fortis, voiced fortis)

(I) consonant	(J) consonant	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
lenis	fortis	-.16384(*)	.01003	.000	-.1888	-.1388
	v_fortis	-.11726(*)	.01410	.000	-.1524	-.0821
fortis	lenis	.16384(*)	.01003	.000	.1388	.1888
	v_fortis	.04659(*)	.01566	.015	.0076	.0856
v_fortis	lenis	.11726(*)	.01410	.000	.0821	.1524
	fortis	-.04659(*)	.01566	.015	-.0856	-.0076

* The mean difference is significant at the .05 level.

Table F: Scheffé's Table for homogeneous subset
Dependent variable: Consonantal Duration
Independent variable: medial consonant type (3 levels: lenis, fortis, voiced fortis)

cons	N	Subset for alpha = .05		
		1	2	3
lenis	57	.0793		
v_fortis	9		.1966	
fortis	21			.2431
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
a Uses Harmonic Mean Sample Size = 17.019.
b The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Table G: ANOVA table
Dependent variable: H1-H2 & H1-A1
Independent variable: medial consonant type (3 levels: lenis, fortis, voiced fortis)

		Sum of Squares	df	Mean Square	F	Sig.
H1-H2	Between Groups	100.592	2	50.296	1.350	.280
	Within Groups	819.674	22	37.258		
	Total	920.266	24			
H1-A1	Between Groups	708.737	2	354.368	3.842	.037
	Within Groups	2029.381	22	92.245		
	Total	2738.118	24			

Table H: Scheffé's Multiple Comparisons Table
Dependent variable: H1-A1
Independent variable: medial consonant type (3 levels: lenis, fortis, voiced fortis)

(I) consonant	(J) consonant	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
lenis	fortis	-.16384(*)	.01003	.000	-.1888	-.1388
	v_fortis	-.11726(*)	.01410	.000	-.1524	-.0821
fortis	lenis	.16384(*)	.01003	.000	.1388	.1888
	v_fortis	.04659(*)	.01566	.015	.0076	.0856
v_fortis	lenis	.11726(*)	.01410	.000	.0821	.1524
	fortis	-.04659(*)	.01566	.015	-.0856	-.0076

* The mean difference is significant at the .05 level.

Table I: Scheffé's Table for homogeneous subset
Dependent variable: H1-A1
Independent variable: medial consonant type (3 levels: voiced fortis, lenis, fortis)

cons	N	Subset for alpha = .05	
		1	2
v_fortis	5	-23.2600	
lenis	13	-19.7462	-19.7462
fortis	7		-9.2429
Sig.		.789	.142

Means for groups in homogeneous subsets are displayed.
a Uses Harmonic Mean Sample Size = 7.147.
b The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.