

Registrogenesis in Bunong: The Interaction of Phonetics and Phonology

Becky Butler

1 Introduction

Registrogenesis is the process by which a voicing contrast on word-initial stops is lost and a resulting contrast remains on following vowels. Register can be phonetically manifested in numerous ways, including pitch, vowel quality and voice quality, and in many languages one of these may serve as the primary register cue while being supported by one or more other cues. However, in Bunong, a Mon-Khmer language of Vietnam, experimental evidence shows that the most salient cues for register are contingent on vowel height. This result suggests that featural attempts to characterize register must be re-examined, and the goal of this paper is to explore a more adequate featural representation of register via the process of phonologization (Hyman 1976).

1.1 Register and Registrogenesis

A register system is one in which a historical voicing contrast in onsets is synchronically manifested on adjacent vowels. Registrogenesis is the process by which this system emerges, when onset voicing is lost and phonological contrast is maintained on the rime. However, it is common that some voicing remain on the stops while other phonetic cues to register begin to emerge on the vowels. For example, Brao, as documented in Huffman (1976) still has a voice ~ voiceless distinction, although the voiced series is slightly aspirated, and vowels following it are “somewhat voiceless and centralized” (579). Nonetheless, phonological register systems can be manifested through phonetics in a variety of ways, including consonantal properties, voice quality, pitch and vowel quality (Brunelle 2005). In most cases, one or two prominent cues are enhanced by other less salient phonetic cues. Examples of common register system differences are given in Table 1. Here and throughout, words with historically voiceless stops are represented by “R1”, and words with historically voiced stops are represented with “R2”.

	First Register (R1) *Voiceless Stops [pa]	Second Register (R2) *Voiced Stops [ba]	Example Languages
Consonantal properties (VOT)	Shorter [pa]	Longer [pha]	Alak (Huffman 1976)
Voice quality/ Phonation	Modal [pa]	Breathy [p̤a]	Mon (Huffman 1976)
Pitch	Higher [pá]	Lower [pà]	Tibetan (Huang 1995)
Vowel quality (F1)	Raised F1/ Lower V [p̚a]	Lowered F1/ Higher V [p̚a]	Khmer (Henderson 1952)
	Plain [pa]	Diphthongized [pia]	Alak (Huffman 1976)

Table 1: Possible Correlates of Register

- (1) Voice Onset Time (VOT): Brunelle (2005), citing Ferlus (1979), suggests that when the voiced stop series is gradually devoiced, voiceless stops become “tense”. One possible result of this tenseness is the development of aspiration on voiceless stops, which then becomes a VOT contrast. Haudricourt (1965) cites three languages – Phay, Samre and Khasi – in which the voiceless series became aspirated stops and the voiced series became plain voiceless stops. Contrastingly, Huffman (1976) notes several languages – Brao, Alak, Souei and Nge? – in which he describes the historically voiced series of stops as “lax and accompanied by slight aspiration” (580). He transcribes these sounds as /p', t', c', k'/, whereas he uses the symbols /p^h, t^h, c^h, k^h/ to transcribe sounds he refers to simply as “aspirated.” The slightly aspirated and (fully) aspirated stops are contrastive in these languages. This suggests that there might be a significant VOT on historically voiced stops, which is distinct from the aspiration found on the series of voiceless aspirated stops. Both of these series contrast with a plain voiceless series.
- (2) Voice quality (phonation): As noted in Table 1, register in Mon is manifested as an R2 ~ R1 contrast between breathy voice and modal voice (Huffman 1976). It is also possible for register to be realized as a contrast between modal voice and creaky voice (Gregerson and Smith 1973). In both scenarios, the vocal folds are held more tightly for R1 vowels than R2 vowels.
- (3) Pitch: Lower pitch is associated with voicing (Kingston and Diehl 1994) and is a common characteristic of vowels following historically voiced stops in register system languages (Huffman 1976, Ferlus 1979, *inter alia*). In many languages, like Vietnamese, for example, pitch distinctions have developed into complete tonal contrast (Matisoff 1973).
- (4) Vowel height: Vowels following voiced sounds often have lower F1s than vowels following voiceless sounds (Lisker 1986). In register systems, this difference can remain on the vowels after the voicing contrast is lost, so that vowels following historically voiced stops are located higher in the vowel space than vowels following historically voiceless stops. This is the case for several register languages, including Khmer, as in (R1) [kom] ‘don’t’ ~ (R2) [kujtiew] ‘noodles’. A change in F1 might not affect the entire vowel, but might only be salient immediately following the onset, causing diphthongization. This is true for Souei (Ferlus 1979), in which F1 is lower at the beginning of many R2 vowels.

1.2 Phonological Background

Because there are so many possible phonetic manifestations of a register system, it is worth asking how any particular cue becomes the primary cue and the basis of contrast. Likewise, it is also necessary to explore the mechanisms underlying the process by which a primary cue loses its perceptual salience, or gets reinterpreted. This pathway is known as phonologization, the process by which intrinsic variation, which may serve as a phonetic cue, becomes phonologically contrastive (Hyman 1976, Ohala 1993).

In order to base the generalizations of Bunong’s register system in some theoretical framework, it is necessary to consider what the minimal unit of contrast in Bunong’s (or any language’s) sound system is. Early conceptions of phonological systems posited that the relevant unit was the distinctive feature, i.e. some abstract unit which minimally differentiated two sounds or classes of sounds. Over time, these features have been called into question (See Cohn 2011 for

a review of the role of features as primitives). Not only has the inventory of features necessary to contrast all sounds in all languages been revised many times, but there have also been inquiries into the acoustic, perceptual or articulatory reality behind these discrete units. In actuality, it has been argued that what binds all the second register acoustic properties together is the lowering of the larynx (Brunelle 2005, Cohn 1993, *inter alia*). However, although each is associated with larynx position, it is only through understanding the phonologization process that we are able to understand why the cues have patterned as they have in Bunong.

1.3 Bunong

Most well-documented cases of register languages are located in Southeast Asia, and are generally Austroasiatic or Austronesian. Although some registrogenesis has been documented in real time, much literature on register systems is reliant upon historical reconstructions of voicing contrasts on initial stops and on synchronic voicing contrasts in related languages (cf. Huffman 1976). This is also true of the present study, in which the voicing is assumed to have existed because it is present in Romanized Bunong orthography, which is used by Bunong speakers in Vietnam, and it is typical of the Mon-Khmer family to which Bunong belongs.

The data presented here were collected from Bunong speakers born in Vietnam who currently reside in the United States. To date, there are very few descriptions of Bunong phonology, with a few exceptions (Bequette 2008, Vogel and Filippi 2006, Phillips 1973). The exact number of phonemic vowels in Bunong is not agreed upon. Bunong may have as many as 18 vowels, including length contrasts for each vowel (Figure 1). However, Phaen et al. 2012 suggest that not all length or height contrasts are contrastive, and Phillips 1973 suggests there are fewer length distinctions than indicated here. However, due to the creation of the register system, the number of vowels has doubled, as will be shown below. The consonant inventory is given in Figure 1. A number of examples of minimal register pairs are given in Figure 2.

	labial	alveolar	palatal	velar	glottal
i					
e					
ε					
ɨ					
ɨ̃					
u					
ʌ					
o					
ɔ					
unaspirated plosive	p	t	c	k	ʔ
aspirated plosive	p ^h	t ^h	c ^h	k ^h	
pre-nasalized plosive	^m b	ⁿ d	^j ɟ	^ŋ g	
implosive	ɓ	ɗ			
nasal	m	n	ɲ	ŋ	
tap		ɾ			
fricative		s			h
approximant	w	l	j		

Figure 1: Bunong Vowels and Consonants

Vowel	Orthography		Vowel	Orthography	
	R1	R2		R1	R2
i	pɨh	bɨh	ʌ	tɔ̃m	dɔ̃m
ĩ/ɨ̃	kĩt	gĩt	u	tu	du
ε	pe	be	ɔ	kop	gop
ɨ̃	tũm	dũm	a	pal	bal

Figure 2: Minimal pairs based on register

As with related languages, Bunong had a voicing contrast on initial stops. This contrast was lost and now Bunong speakers use various cues to maintain contrast due to the loss of voicing on initial stops, which are explored in the experiments below. These cues range from vowel quality to voice quality to differences in consonantal properties, and sometimes interact with each other in unexpected ways. Furthermore, because Bunong is still in the process of registrogenesis, it provides a wealth of information about how registrogenesis occurs, which might be more difficult to see in a language like Khmer (Wayland and Jongman 2003), in which the register system is well established.

Bunong register is remarkable for at least two reasons. First, the phonetic manifestation of register is contingent on the phonology of the language, such that the observed differences are dependent on vowel height. Specifically, diphthongization is a cue to R1 on high and mid vowels but is a cue to R2 on low vowels. Additionally, pitch and phonation contrasts are dependent on vowel height, which suggests an interaction between the phonology and the phonetics. Second, in Bunong, pitch is higher following historically voiced stops than historically voiceless stops, contrary to what is expected, given the acoustic correlates of larynx positioning.

In order to understand these issues and to answer the theoretical questions raised at the end of §1.1, I present two phonetic experiments on Bunong. The first is a production experiment which aims to provide a general understanding of the Bunong vowel system and to investigate phonetic correlates of register in Bunong. The second is a perception experiment which aims to shed light on the perceptual saliency of phonetic cues for register.

2 Experiment 1: Production

The goal of this production experiment is to assess which phonetic cues Bunong speakers use in their maintenance of lexical contrast through phonological register. Using vowel tokens with three different onsets from both registers, five possible cues to register were measured on the vowels: HNR, H1-H2, F1 Bandwidth, VOT and Pitch. Statistical analyses show that each cue is significantly correlated with register. Furthermore, formant values show that there are consistent differences in vowel quality between registers.

2.1 Methodology

To determine which phonetic properties Bunong speakers use to differentiate vowels, recordings were made of a wordlist comprising ten vowels and six initial stop consonants p_{R1} ('p'), t_{R1} ('t'), k_{R1} ('k'), p_{R2} ('b'), t_{R2} ('d'), and k_{R2} ('g'). Each consonant was followed by ten different vowels, for a total of 60 stimuli (See Appendix 1 for wordlist), and three to four repetitions of each were recorded, resulting in minimally 180 tokens per speaker. Words were recorded in lists, without the use of a frame sentence. Tokens were actual words of Bunong, so in some cases, it was not possible to control for codas across the entire data set, but efforts were made to control the codas in minimal pairs. For example, ' $t_{R1}ak$ ' and ' $t_{R2}ak$ ' (Figure 1 and Figure 2, respectively) have the same coda, so that direct comparisons can be made between them. Notice that in both spectrograms, the onset is entirely voiceless, but that the vowel formants are quite different.

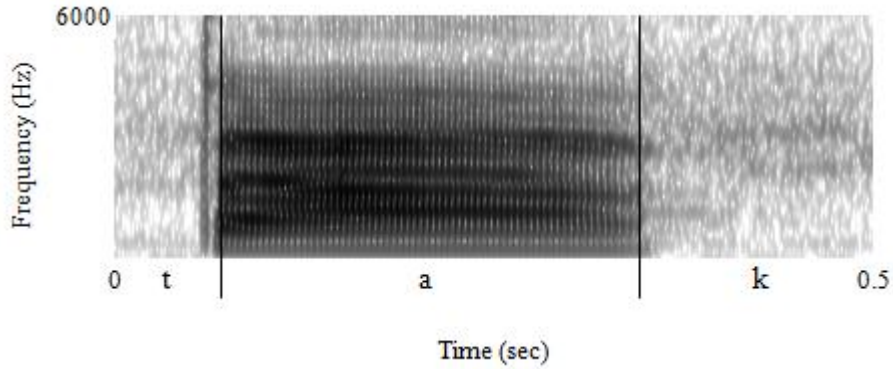


Figure 1: Spectrogram of *tak* (R1) [tak]

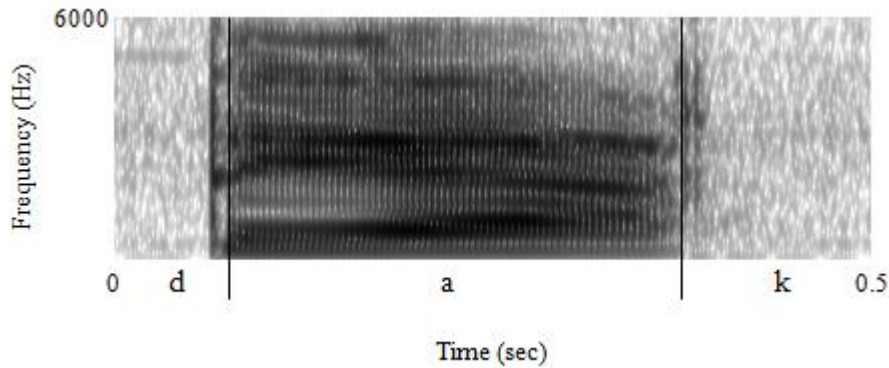


Figure 2: Spectrogram of *dak* (R2) [teak]

Four Bunong speakers were recorded. All speakers were from Vietnam, and all were in their late 20s/early 30s. All had immigrated to the US within the last seven years at the time of recording. Two of the speakers were men (YS, YK) and two were women (HB, HN). Recordings were made using a Logitech headset microphone (model #A-0356A) and a laptop computer, and measurements were made in PRAAT (Boersma and Weenink 2007). The four cues to register – consonantal properties, voice quality, pitch and vowel quality – were measured in turn.

(1) Consonantal properties

These are used to determine whether register cues are realized exclusively on following vowels or if the contrast is still in part associated with the consonant.

- a) VOT was measured from the beginning of the stop burst to the onset of F2 in the following vowel.
- b) Voicing during the stop was determined by whether or not glottal pulses were visible.

(2) Voice quality

Phonation can have several acoustic correlates. The following three measures were used for this study. Based on impressionistic data, Bunong seems to have a breathy ~ modal contrast but not a modal ~ creaky contrast, and measurements are predicted to suggest less vocal fold tension for R2 vowels than for R1 vowels.

a) $\Delta H1-H2$

The difference in amplitude between the first harmonic (H1) and the second harmonic (H2) can be an indicator of breathiness. Because breathy vowels have a longer open

quotient than modal vowels, their spectral shape is more similar to a sine wave than is the spectral shape of modal or creaky vowels (Johnson 1997). Thus, the prediction is that breathy vowels will have a larger $\Delta H1 - H2$ than modal vowels. Measurements were made via a script retrieving the maxima for the first and second harmonics of L_{tas} spectra.

b) F1 Bandwidth

Breathy vowels are claimed to have a larger first formant (F1) bandwidth value than modal vowels (Kent & Read 2002). The size of the F1 bandwidth is related to the amount of absorption of sound energy, i.e. damping. The greater the damping, the larger the bandwidth. Breathless vowels allow for more damping than modal vowels because of the looser position of the vocal folds during phonation. This measurement was taken using the query function in a Burg object at the first quarter point.

c) HNR

Harmonics-to-noise ratio is the comparison of periodic noise to aperiodic noise in a signal (Boersma 2003). Breathless vowels have more noise than modal vowels because the vocal folds are not held as tightly, so there is a longer open period during which there is a high rate of airflow. Therefore, breathless vowels should have a lower HNR than modal vowels. HNR measurements were taken from a Praat Harmonicity object (Boersma and Weenink 2007) via the query function at the first quarter point.

(3) Pitch

In the recorded data, pitch is generally held steady over the duration of the vowels, with the exception of a slight downtrend. Although repetitions were collected in a list, there was no appearance of list effects. Because pitch contours were found to be relatively flat, measurements were taken at the quarter point and halfway point and then averaged together. Because lower pitch is known to correlate with consonant voicing, R2 vowels are predicted to have a lower pitch than R1 vowels. Pitch measurements were taken using Fast Fourier Transform spectra. Additionally, we expect there to be an effect of gender on pitch, since women generally have higher fundamental frequencies than men.

(4) Vowel quality

F1 and F2 were measured at 25% and 50% of the duration of the vowel. F1 is predicted to be higher for R2 vowels than for R1 vowels. This may either be a consistent height difference throughout the duration of the vowels, or we may see diphthongization created by changes in F1 very near the consonant but not later on in the production of the vowel. The second formant is not expected to have any effect due to register, but it is useful for determining where vowels are located in the vowel space.

2.2 Results

Three of the four possible predictors of register – consonantal properties, voice quality and pitch – were analyzed in a mixed model (Vowel quality was tested separately). Dependent variables included VOT, HNR, $\Delta H1-H2$, F1 Bandwidth and Pitch. Independent variables were Register (R1 or R2), Gender (M or F) and the interaction of Register with Gender, as well as Speaker, which was a random variable. Tests were run by each vowel height – high, mid and low. Significance levels for each correlation are presented in Table 3 below. Lack of significance is indicated with an NS (not significant).

Vowel Height	Factors				
	VOT	HNR	H1-H2	F1 Bandwidth	Pitch
High	NS	NS	0.0013	0.0128	0.0001
Mid	NS	0.0002	0.0476	NS	<0.0001
Low	NS	<0.0001	NS	0.0004	0.0004

Table 3: Significance of correlation of factors with the interaction of Register and Gender, by vowel height

(1) Consonantal Properties

VOT is highly correlated with register at all vowel heights ($p < 0.001$). In other words, R2 vowels have significantly longer VOT than R1 vowels. VOT is not correlated with gender and as shown in Table 3, VOT is also not significantly correlated with the interaction of register and gender for any vowel height. In addition, for all tokens, it was confirmed visually on the spectrogram that no glottal pulses are present in the onsets.

(2) Voice Quality

HNR for high vowels, F1 Bandwidth for mid vowels and H1-H2 for the low vowel are significantly correlated with register alone suggesting that gender does not make a difference in these cases. In most cases, however, measures of phonation difference (HNR, H1-H2 and F1 Bandwidth) are significantly correlated with the interaction of register and gender at all vowel heights.

Because Δ H1-H2 for mid vowels is marginally significant at a level of $p = 0.05$, an additional ANOVA was run on individual speakers for this factor. The results (Table 4) show H1-H2 to be highly correlated with register in mid vowels for speaker YK, somewhat correlated for HB, yet not at all correlated for speakers HN or YS. However, when HNR, H1-H2 and F1 Bandwidth are considered together, the overall trend indicates that register is correlated with phonation type, such that R1 vowels are modal and R2 vowels are breathy.

Speaker	YK	HB	HN	YS
Level of significance	<.0001	0.025	NS	NS

Table 4: ANOVA results by speaker for H1-H2 of mid vowels

(3) Pitch

For high and mid vowels, pitch was higher for R2 vowels than for R1 vowels. Since voicing is usually correlated with a lowered pitch (Kingston and Diehl 1994), it is surprising that this pattern has emerged. One possible explanation is that since the F1 of R2 vowels decreased, at least vowel-initially, the associated intrinsic pitch may have simultaneously risen (cf. Whalen and Levitt 1995). However, pitch is almost identical on low vowels following both historically voiced and historically voiceless stops. Nonetheless, pitch is significantly correlated with register, and the interaction of register and gender is significantly correlated with pitch at all vowel heights ($p = 0.05$).

(4) Vowel Quality

A representative vowel space (i.e. F1 by F2) for one speaker is shown in Figure 3. It contains the vowels /i, e, u, o, a/, which were chosen as a representative sample of vowel heights.

Each vowel has a beginning point and an end point, the former represents the formant measurements taken at 25% of the duration of the vowel, and the latter represents the measurements taken at 50% of the duration of the vowel. Plotting two points for each vowel allows us to see not only the extent of diphthongization but also the direction of the movement in the vowel space. Measurements were taken from all tokens containing these vowels, including words with labial, alveolar and velar onsets, and all measurements for each vowel were averaged together to avoid skewing that might have been caused by the place of articulation of the preceding consonant.

There are several salient characteristics of the representation of speaker YK’s vowel space below, which are common to the vowel spaces of the other speakers as well. First, R2 vowels are consistently higher in the vowel space (lower F1) than the equivalent R1 vowels. Additionally, high and mid R1 vowels are much more diphthongal than their R2 counterparts, as evidenced by the lengths of the lines connecting the beginning and end points of each vowel. However, the opposite is true for the low vowel /a/, in which the R2 vowel is much more diphthongal than the R1 vowel. Furthermore, for most vowels, diphthongization proceeds toward the periphery of the vowel space.

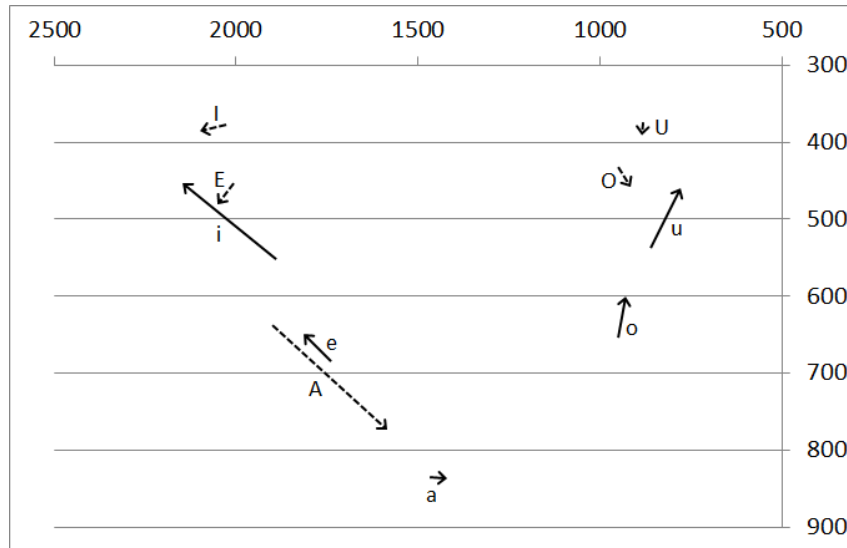


Figure 3: Vowel Space for Speaker YK, showing R1 and R2 /i,e,a,o,u/ averaged over POA. R1 is represented by solid lines and lowercase letters. R2 is represented by dashed lines and capital letters.

Based on the F1 and F2 measurements, the degree of diphthongization is significantly greater ($p = 0.05$) for high and mid R1 vowels than R2 vowels, respectively, in Figure 4. This was measured by determining the Euclidean distance between the x,y coordinates taken at the beginning of the vowel and the x,y coordinates taken at the halfway point of the vowel. In sum, high and mid R1 vowels are more diphthongized than high and mid R2 vowels. In contrast, the low vowel [a] shows the opposite pattern, such that R2 vowels are more diphthongized than R1 vowels.

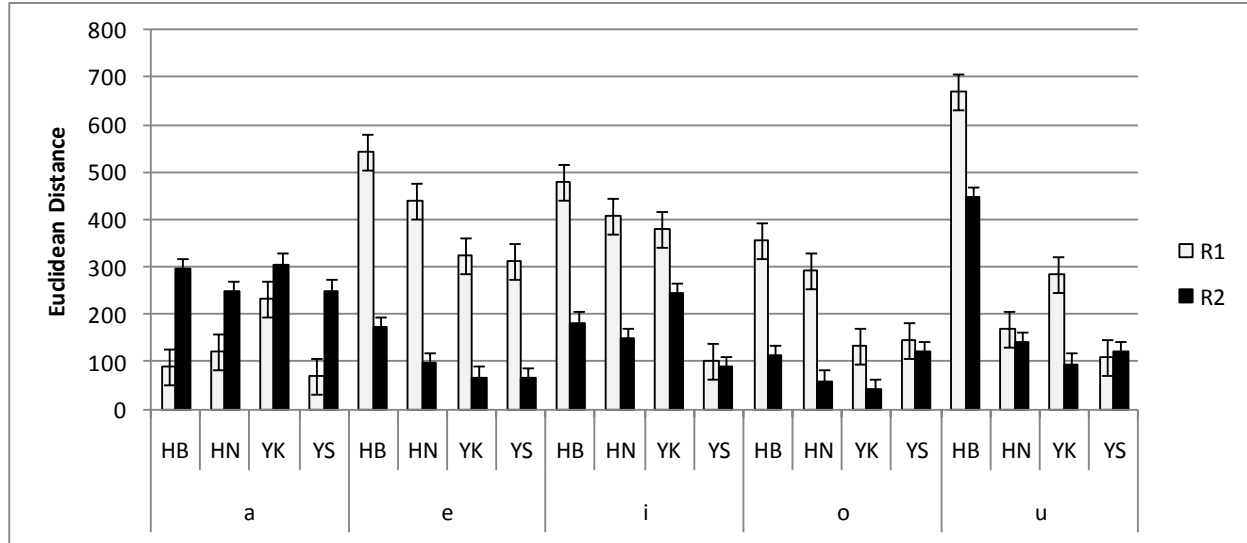


Figure 4: Degree of diphthongization of vowels (Hz), by speaker

To summarize, although the degree of diphthongization does vary across speakers, the overall pattern shows that non-low R1 vowels are diphthongs and low R2 vowels are diphthongs. These findings are supported by patterns of diphthongization, which is longer for R1 non-low vowels than for R1 low vowels and for R2 low vowels than for R2 non-low vowels. When averaged across speakers, these differences are significant ($p = 0.05$).

2.3 Discussion

In general, results show no remaining voicing contrast on initial consonants. Instead, consonantal properties and voice quality are correlated with register in expected ways. VOT is correlated with register alone, and HNR, H1-H2 and F1 Bandwidth are also generally correlated with register when gender is accounted for. While pitch and vowel quality also correlate with register, the direction of the correlation is somewhat unexpected. In particular, for non-low vowels, pitch is higher following historically voiced stops and there is a greater degree of diphthongization following historically voiceless stops.

Generally, when diphthongization occurs as a result of a loss of voicing, the vowels following historically voiced stops diphthongize (as is the case with the low vowel in Bunong), not the vowels following historically voiceless stops. However, if diphthongization is seen as a partial change in F1, i.e. an F1 change at the beginning of the vowel, which is most likely to be affected by registrogenesis due to its proximity to the onset, diphthongization of R1 vowels is less surprising. In contrast to the low vowel R2 diphthong, non-low vowel R1 diphthongs are all rising. This indicates that an increase in F1 occurred at the beginning of R1 vowels.

Although the pattern of diphthongization is phonetically motivated, it is noteworthy that this F1 increase occurred only in non-low vowels and the F1 decrease occurred only in low vowels, instead of one or the other process happening across all vowel heights. This suggests, first of all, that vowel height is having an effect on the phonetic manifestation of register. Second, as can be seen in Figure 3, the F1 of mid R2 vowels decreased and the F1 of high R1 vowels increased, so that they overlap in the present day vowel space. Because their vowel qualities are now so similar, diphthongization is likely used to enhance the difference in the vowels.

A similar phenomenon has been noted in New Zealand English (NZE) by Maclagan and Hay (2007). They suggest that because [ɛ] has raised to the height of [i], [i] has diphthongized in order to preserve the contrast between the two vowels. They found that the smaller the Euclidean Distance between [ɛ] and [i], the greater the diphthongization of [i], and that these two measures are significantly correlated. The above results suggest that in Bunong, just as in NZE, phonetics are manipulated in order to maintain phonological contrast.

3 Experiment 2: Perception

The results of Experiment 1 suggest that cues employed by Bunong speakers vary across vowel heights and between speakers. The following experiment is a first attempt to determine which cues are most closely attended to by listeners and how these cues are interpreted. This experiment is a forced choice task, in which listeners are presented with two stimuli and must determine which of two possible words they are hearing. For each token, participants also provide a goodness rating from 1 to 7 in order to indicate how confident they are in their choice.

3.1 Methodology

Stimuli

Stimuli were created from natural speech previously recorded from a male Bunong speaker. Because effects were predicted to vary by the height of the vowel, stimuli were used for three phonological heights – high, mid and low – given in Table 5 below. At each height, words were minimal pairs. Each minimal pair was cross-spliced, once with the historically voiced onset preceding the vowel following the historically voiceless onset [C_{R2}V_{R1}] and once vice versa [C_{R1}V_{R2}]. For each new cross-spliced word, four tokens were used. One token was left natural, i.e. no changes were made to VOT or pitch, but the token was still cross-spliced. In another token, the VOT was altered; in another, the pitch was altered; and finally, in the fourth token both the VOT and pitch were altered.

Vowel Height	Transcription/Gloss	
	R1	R2
high	tih 'sin'	tih 'incest'
mid	təm 'plant'	təm 'forgiveness'
low	tak 'sheath'	tak 'water'

Table 5: Minimal pairs used in Experiment 2

Voice Onset Time

For the tokens from the speaker on whom the stimuli were based, VOT was longer for R2 stops than R1 stops. VOT was an average of 14ms, 13ms and 10ms longer for high, mid and low vowels, respectively. When VOT values were changed for a given token, they were either reduced or elongated to match the pattern of the stop with the opposite historical voicing.

Pitch

To alter the non-low stimuli, pitch on vowels following R2 stops was lowered 10Hz, and pitch on vowels following R1 stops was raised 10Hz. However, because pitch in the natural data was so similar on both R1 and R2 low vowels, the pitch for the stimuli was both raised and lowered 10Hz following R2 and R1 stops. Therefore, there were more stimuli for the low vowels than for the high or mid vowels.

Task

The experiment comprised four Bunong participants – two male and two female; two of whom also participated in the acoustic study. Both males and one female were in their early 30's, and the second female was in her late teens. All speakers were literate in Bunong. Stimuli were separated by vowel height, so that they could be presented as minimal pairs, and the stimuli in each group were presented in random order. Participants were given a choice of two words, e.g. *tih* and *dih*, and asked to write which they heard. For every token, they were asked to provide a goodness rating from 1 to 7, to indicate how natural each token sounded. They were allowed to hear the stimuli as many times as needed, though twice was usually sufficient.

There were several possible outcomes to this experiment. If speakers chose to place the tokens in a register based on one cue only, even when other cues were conflicting, then that cue can be considered the primary cue for register. However, it was also possible that speakers would chose to categorize stimuli when two cues were working together contra the third cue. These may have always been the same two cues, or it might have been the case that speakers would decide based on the majority of cues, whatever those cues might be. In the latter case, it is likely that none of the cues is more important than the other; all are playing an important role. If speakers divided the stimuli based on vowels only, without regard to pitch or VOT, then it would be possible to conclude that either (i) position within the vowel space overrides pitch and VOT or that (ii) phonation type overrides pitch and VOT; however, it will not be possible to determine which characteristic of the vowel is most salient. At the very least, this experiment will show where the contrast lies – (i) in the consonant (VOT), (ii) in the vowel quality (formants) or voice quality (phonation) or (iii) in the pitch.

3.2 Results

The forced choice task and the goodness ratings show that participants (P1, P2, P3 and P4) generally determine register based on cues in the vowel, either voice quality or vowel quality, rather than pitch or VOT. Nonetheless, results show that pitch and VOT still have effects on perceptions of register, which is particularly evident in the goodness ratings. This is true across vowel heights, though notably two of the participants are less sensitive to the pitch and VOT manipulations on low vowels than on non-low vowels (See Appendix 2 for all results).

3.2.1 High vowels

For high vowels, participants overwhelmingly chose R2 when the formants of the vowel were R2 (indicated by the black bars in Figure 5) and R1 when the formants of the vowel were R1 (indicated by the light gray bars), despite alterations in VOT or pitch. For one token in which the pitch on a V_{R2} was altered to be R1, P2 chose R1 instead of R2. However, P1's choices were largely based on the original register of the onset, but were nonetheless sensitive to alterations in pitch.

However, goodness ratings, (y-axis) suggest there may be more sensitivity to changes in pitch and VOT than the forced choice task reveals. P4, for example, gives slightly lower goodness ratings when the R2 vowels have R1 pitch values. P1 and P2 give consistently lower goodness ratings than P3 and P4 in general. Interestingly, P1 gives lower goodness ratings for all $C_{R_2}V_{R_1}$ sequences. These results suggest that although P4 shows some sensitivity to pitch and VOT for high vowels, P1 and P2 are more strongly attuned to cues other than formant values than are P3 and P4.

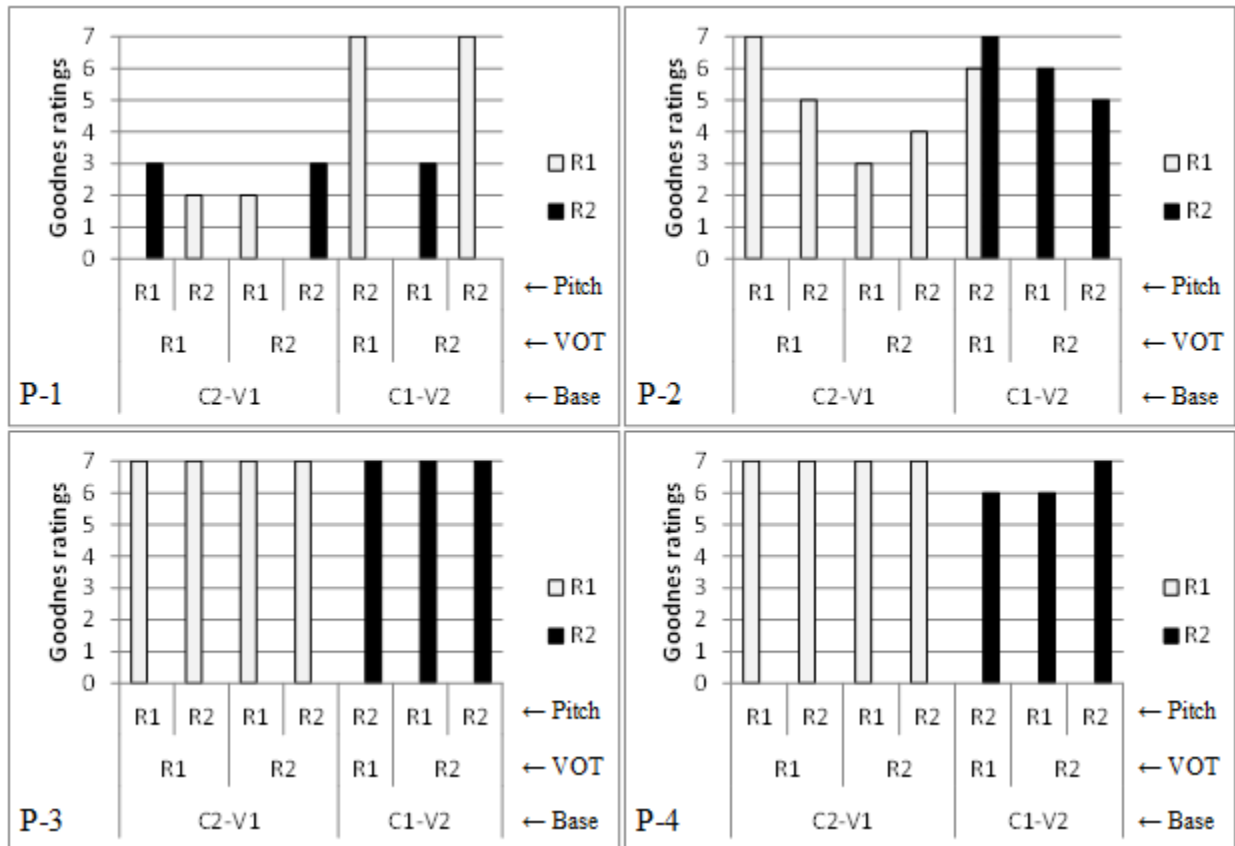


Figure 5: Goodness ratings for forced choice task of cross-spliced high vowel tokens

3.2.2 Mid vowels

Forced choice results are less variable for mid vowels than for high vowels. P1, P3 and P4 chose only according to the original register specification on the vowel, indicating that they are highly attuned to formant and phonation values. P2, however, largely chose according to onset. For $[C_{R_1}V_{R_2}]$ tokens, P2 chose in accordance with the VOT, such that if the VOT was altered to be R2, P2 chose R2. For $[C_{R_2}V_{R_1}]$ tokens, P2 seems to have chosen mostly based on the original register specification of the onset, except in one case where R1 is chosen despite the token having R2 VOT and pitch.

Nonetheless, goodness ratings for mid vowel stimuli are lower on average than they are for high vowel stimuli. P3 and P4 consistently gave lower goodness ratings to $[C_{R_2}V_{R_1}]$ sequences than to $[C_{R_1}V_{R_2}]$ sequences, regardless of pitch or VOT. Interestingly, although P2's judgments in the choice task were the most variable, P1 gave overall lower goodness ratings than P2, particularly when the pitch did not match the formant values of the vowel. Just as the results

for high vowels, the goodness ratings for mid vowels suggest that although register choice seems primarily based on formant values, speakers are still attuned to changes in pitch and VOT.

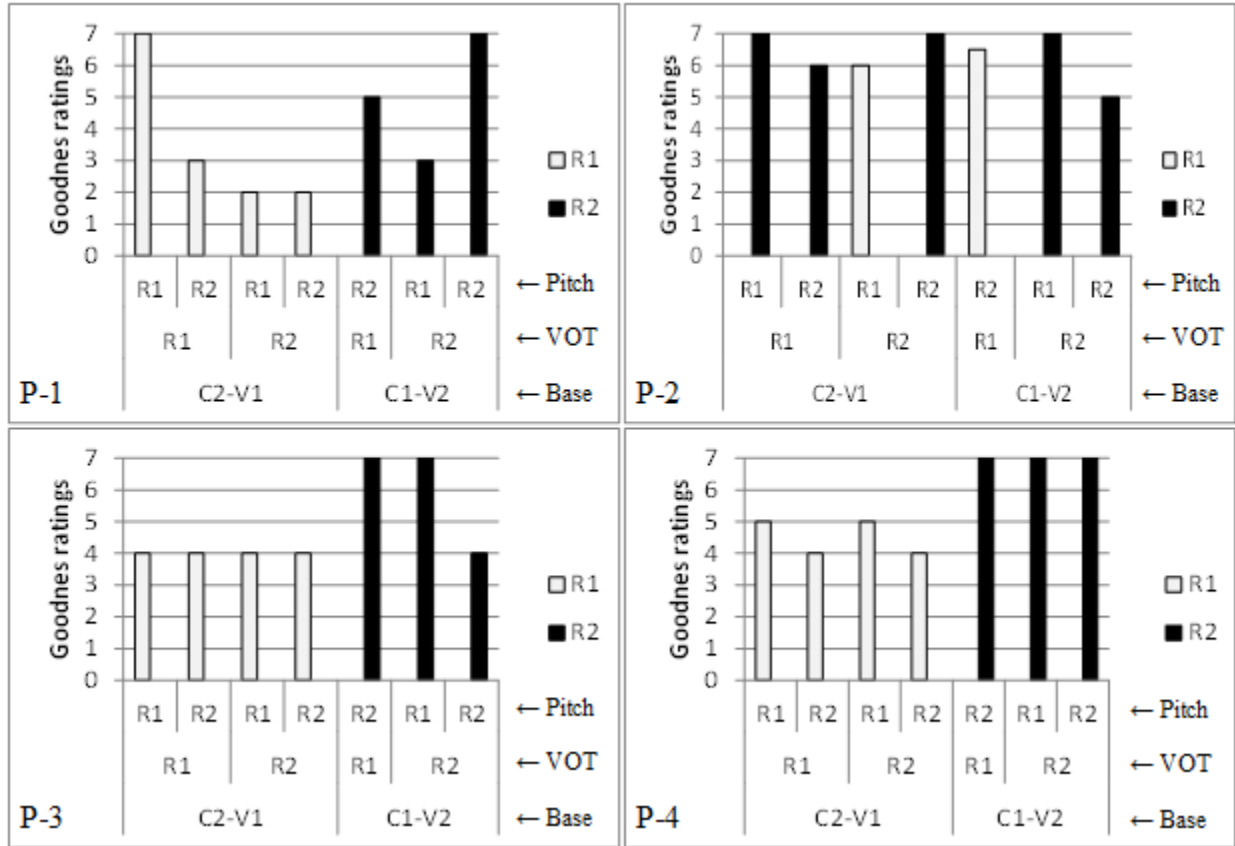


Figure 6: Goodness ratings for forced choice task of cross-spliced mid vowel tokens

3.2.3 Low vowels

For low vowels, participants again chose register class based largely on the formants and phonation of the vowel, regardless of pitch changes or VOT alterations on the preceding consonant. There were three cases in which P2 chose differently. First, in two cases, when the consonant had an R2 VOT and the pitch was altered on the R1 vowel, P2 chose R2. It is surprising that both a raising and a lowering of pitch would have the same effect, but P2's goodness ratings for both of these choices are relatively low. In addition, P2 chose R1 for one token with an R2 vowel but an R1 consonant. In all three of these cases, R2's choice is aligns with the register specification of the VOT.

Of particular interest regarding the low vowel is that P3 and P4 give consistently perfect goodness ratings despite changes in pitch or VOT. This suggests that in the case of this vowel, vowel quality is such a strong cue, alternations in other cues are not perceptually salient enough to affect register judgments. P1 gives higher goodness ratings for R2 vowels with higher pitch and R1 vowels with lower pitch, regardless of alterations on VOT, but P2 gives the highest goodness ratings when there are no pitch alterations. P2 seems to give generally higher goodness ratings to stimuli with R2 vowels, no matter the pitch or VOT.

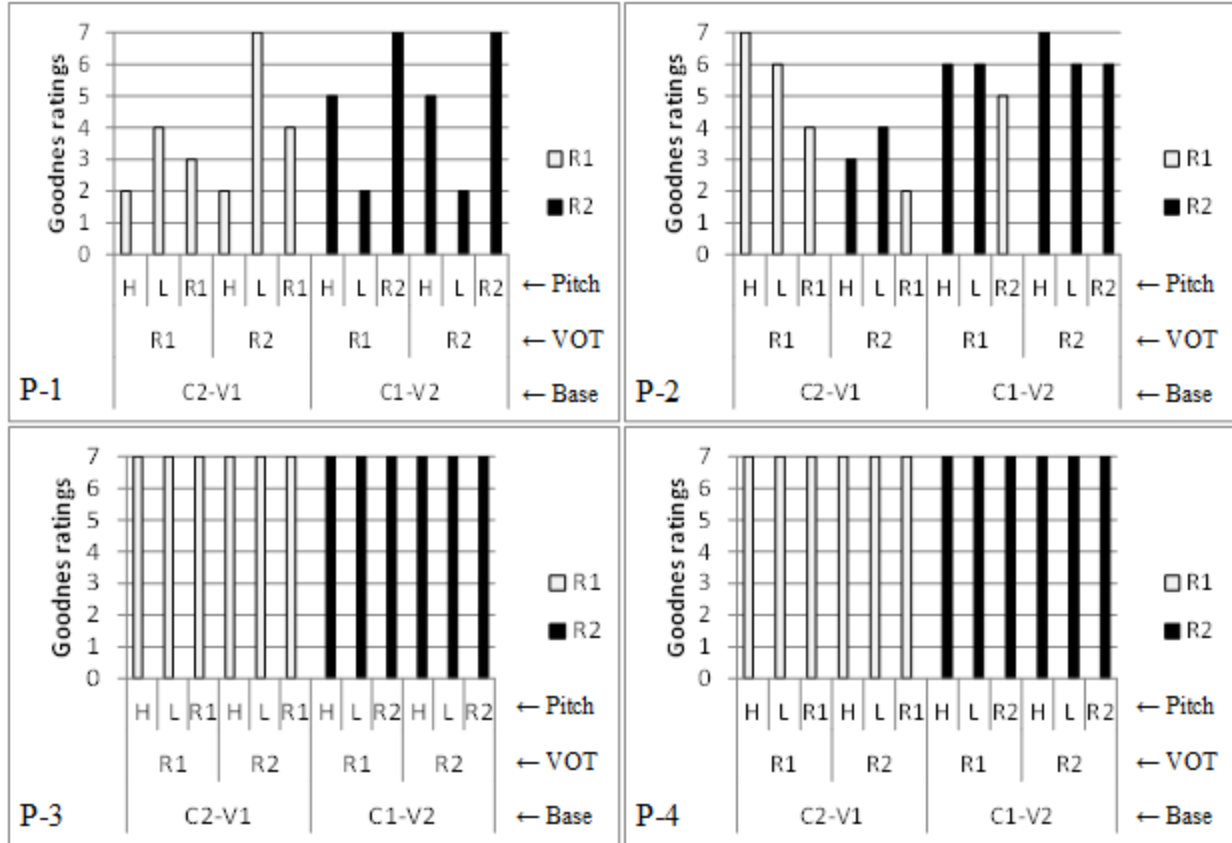


Figure 7: Goodness ratings for forced choice task of cross-spliced low vowel tokens

3.3 Discussion

The above results suggest that the most salient cue to register is vowel quality (diphthongization or vowel height) or phonation. However, although participants largely chose register based on properties of the vowel regardless of VOT and pitch, goodness ratings suggest that these two cues are still playing a role. Participants often gave lower goodness ratings when the register of the pitch did not match the register of the vowel formants.

Additionally, there are differences between responses given for low and non-low vowels. Notably, judgments were far more consistent on the low vowel stimuli than the high and mid vowel stimuli. This suggests that Bunong speakers may be more attuned to changes in pitch or VOT in non-low vowels than in low vowels. Therefore, phonological register likely encompasses various phonetic manifestations, dependent on the phonological height of vowels.

4 General Discussion

In light of the above experiments, I now turn to the question of the phonological status of register. In accordance with Repp (1986), who suggests that a distinctive feature may comprise more than one phonetic cue, i.e. there is not necessarily a one-to-one correlation between phonological category and phonetic realization, the results suggest that the feature [register] represents a bundle of phonetic correlates discussed in the experiments above, which can vary over time. In addition, viewing registrogenesis as a type of phonologization (Hyman 1976) provides an understanding of how the feature system changes over time and how this asymmetric relationship between phonology and phonetics may arise.

4.1 *The Role of Cues*

Stevens et al.'s (1986) Enhancement Theory posits that some “phonetic feature” works as a primary cue in a contrastive system, and this cue is enhanced by other non-arbitrary cues. All cues, whether primary or enhancing, are taken from a finite list of universal “distinctive features.” For example, rounding is said to enhance backness because of “the low-frequency spectral prominence resulting from the proximity of F1 and F2” (442). They further suggest that it is possible for primary and enhancement cues to switch places, such that some enhancement cue can become the primary cue. Enhancement is necessary only when there might otherwise be confusion between two primary cues; therefore, enhancement cues are used to increase perceptual distance.

This perspective has been shown to be problematic, for as Repp (1986) states when commenting on Stevens et al (1986), there are more cues in the phonetic representation than there are in the distinctive feature inventory. Indeed, this view largely oversimplifies the subtleties of the phonetic system. Additionally, the non-arbitrariness of cues is problematic. For example, although breathiness and high pitch are correlated with each other in the Bunong register system, they do not enhance one another phonetically.

More recently, Kingston and Diehl (1994) recast enhancement cues in terms of covariation of acoustic cues and their articulations. They maintain that speakers covary cues intentionally and that speakers are aware of the consequences their articulations have on the phonetic output. For example, because no articulation other than voicing affects f_0 (voiced stops correlate with a lower f_0 than voiceless stops), then a depression of f_0 “must be the product of an independently controlled articulation, whose purpose is to enhance the [voice] contrast” (425). However, despite intentionality on the part of speakers, the authors claim that phonetic manipulation for the sake of redundancy [or enhancement] is not necessarily phonologized. Indeed, this is what separates the covariation theory from the Enhancement Theory, namely that in the latter, redundancy can exist in the phonetic component of speech, not just in the phonology. Subsequently via experiments by Kingston and Macmillan (1995) and Turk and Sawusch (1996), this line of thought has been expanded into considerations of cue trading.

More recent work focuses on the probabilities of cue trading and cue weighting. Clayards (2008) investigated cue trading in terms of likelihood distributions of acoustic phonetic cues. The model she employs qualitatively predicts cue weights, and she suggests that speakers make use of the variation in acoustic cues they receive in the speech signal. Because this approach focuses on how variation can be used to convey information, it does not provide an analysis of the types of discrete categories found in the phonology. Because register in Bunong displays variation in both production and perception and across speakers, this may be seen as evidence for a probabilistic model approach. However, in the perception experiment, variation in the signal only served to confuse participants. Instead, speakers seem to use various cues as a bundle for a register feature.

Thus far, much of this work has been focused on rethinking Jakobsonian type features. However, symbolic feature systems that connect phonetics to phonology have been shown to be problematic in essence, i.e. they require some mechanism for mapping a continuous reality to a discrete representation. This is indeed problematic for describing emergent register systems because the vast range of phonetic possibilities of the feature [register] are seemingly quite unrelated to each other and to the [voice] feature. However, larynx lowering is thought to be related to all the manifestations of R2, including breathiness, voicing, lowered pitch, etc. Indeed,

Results from Experiment 1 show that the vowel height ~ stop voicing relationship was enhanced, such that vowels following historically voiced stops raised. This presumably occurred in the extrinsic (phonological) stage. As a result of this enhancement, vowel height was then used to cue the voicing feature instead of voicing cueing vowel height. This change in F1 likely also accounts for the diphthongization of the low vowel. The height of the vowel simply rose more at its beginning, directly adjacent to the voiced stop, than at its end, creating a diphthong.

At this point, it is surprising that this process did not also apply to high and mid vowels. Instead the F1 contrast was enhanced on vowels following voiceless stops, such that their F1 was raised dramatically following voiceless stops, thereby turning them into diphthongs, as well. Although the changes in F1 follow directly from the intrinsic properties of vowel height and voicing, it is not at all clear why some of these changes were enhanced on vowels following voiceless stops and some on vowels following voiced stops, instead of one set alone.

In addition to the vowel height changes, the association of vowel height and pitch was also enhanced at this second stage. Although register systems involving pitch usually show the opposite (and more transparent) pattern, where pitch is lowered following voiced stops, the opposite is true in Bunong. This is likely because the change in pitch occurred after or concurrent with the vowel height changes. Because vowels following voiced stops were raised, pitch on non-low vowels following voiced stops was also raised in order to enhance the vowel height difference. It is possible that this intrinsic feature would not have passed from the intrinsic to the extrinsic stage if the vowel height contrast had not previously done so.

Step 1 $F1 \rightarrow \downarrow / [+voi]$
Voicing raises R2 vowels

$V_{+low} \rightarrow V_{diph}/C_{+voi}____$
Low vowel diphthongizes following voiced stops due to the F1 lowering

Step 2 $F1 \rightarrow \downarrow / [+voi] \Rightarrow [\text{register}] \rightarrow [R2] / F1_{\downarrow}, [\text{register}] \rightarrow [R1] / F1_{\uparrow}$
Voicing ceases to determine vowel quality. Instead, vowel quality determines register.

Non-low vowel categories conflate

Due to conflation of non-low vowel categories:

$V_{-low} \rightarrow V_{diph}/C_{-voi}____$
Non-low vowels diphthongize following voiceless stops

Step 3 $f0 \rightarrow [\uparrow]/V_{-low\uparrow}$
Pitch contrast is enhanced on non-low vowels, so higher vowels have a higher pitch

$V_{-low} \rightarrow \underset{\cdot}{V}/C_{+voi}____$
Phonation contrast emerges on non-low vowels

Result $[\text{register}] \rightarrow [2]/F1_{\downarrow}, f0_{\uparrow}, \underset{\cdot}{V}, [\text{register}] \rightarrow [1]/F1_{\uparrow}, f0_{\downarrow}, V_{\text{modal}}$
Vowel quality, pitch and phonation determine register on non-low vowels

Figure 8: Development of phonological stage

Although the first two stages can be explained in these terms, the third (phonemic) stage is less obvious. The phonemicization process requires that, for example, the use of $f0$ or $F1$ to cue the voicing feature be dropped entirely and instead that pitch or vowel height constitute the phonemic contrast while being augmented by other cues. However, in Bunong it is not the case that any one cue holds this status exclusively. VOT, pitch, phonation and vowel height all work in conjunction to cue the feature [register]. Although voicing has been eliminated entirely on onsets, the register feature remains and represents a bundle of phonetic cues which cannot be determined entirely by vowel height.

5 Conclusion

What does this say about the representation of cues in Bunong? In simplest terms, the phonetics and phonology are dependent on one another. In addition, these data speak to the amount of variability in phonological representations. Keating (1984) suggests that although the

phonetic realization of a phonological feature may vary across languages, restrictions on phonetic systems ensure that such variation is non-arbitrary. For example, in a study of Polish, English and German speakers, she finds that the contrast [+voice] ~ [-voice] is realized as contrast between voiced stops ~ voiceless unaspirated stops or as a contrast between voiceless unaspirated stops ~ voiceless aspirated stops. What she does not find is a contrast between voiced stops and voiceless aspirated stops, suggesting that if a language only has one type of stop category, it will be voiceless unaspirated.

The Bunong data, however, produced a more varied system. First, phonation contrasts generally follow a type of non-arbitrary contrast similar to stop voicing systems, where R2 is breathy and R1 is modal or else R2 is modal and R1 is creaky. There do not appear to be any cases in which a language contrasts breathy and creaky phonation types without also having modal phonation. In this sense, Bunong is quite typical in having a modal ~ breathy contrast. Next, however, the case of diphthongization is more complicated. Although the change in F1 at the beginning of the vowel in Bunong is predictable by register, diphthongization itself is not, since R1 non-low vowels diphthongize but R2 low vowels diphthongize. Finally, pitch is particularly contrary, since it behaves in an opposite way than predicted. Although the pitch pattern can be reasonably explained by vowel height changes, its correlation with register is not transparent.

In summary, although the data can be accounted for, the change from a voicing system to a register system in Bunong produces some unexpected results. Nonetheless, what ties all the phonetic cues together is speculatively the position of the larynx. This suggests that the data may be more adequately accounted for in an articulatory model, which uses gestures – like laryngeal position – as primitives. Whatever the theory, it is clear that various acoustic outputs, most likely resulting from laryngeal movement, represent identical cognitive units to Bunong speakers since each can be an indicator of register. This suggests that what results from larynx position at a gestural level is represented on a more abstract phonological level as the feature [register], which varies both across and within speakers.

References

- Bequette, Rebecca. 2008. *Participant Reference, Deixis, and Anaphora in Bunong Narrative Discourse*. Master's Thesis. Graduate Institute of Applied Linguistics.
- Boersma, Paul. 2003. PRAAT documentation.
- Boersma, Paul & Weenink, David. 2007. *Praat: doing phonetics by computer*. Version 4.6.40 [Computer program]
- Browman, C., & L. Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook*, 3:219-252.
- Brunelle, Marc. 2005. Register in Eastern Cham: Phonetic, Phonological and Sociolinguistic Approaches. PhD Dissertation. Cornell University.
- Clayards, Meghan. 2008. The Ideal Listener: Making Optimal Use of Acoustic Phonetic Cues for Word Recognition. PhD dissertation.
- Chitoran, Ionan and Cohn, Abigail. 2009. Phonetics and phonology: gradience, categoriality and naturalness. *Approaches to Phonological Complexity*, ed. by F. Pellegrino, E. Marisco, and C. Coupé, 21-46. Berlin: Mouton de Gruyter.
- Cohn, Abigail. 2011. Features, segments and the sources of phonological primitives. *Where Do Features Come From?*, ed. by G. N. Clements and R. Ridouane, 15-41. Amsterdam: John Benjamins.
- Cohn, Abigail. 1993. Consonant-vowel interaction in Madurese: The feature Lowered Larynx. *CLS* 29:105-119.
- Ferlus, Michel. 1979. Formation des registres et mutations consonantiques dans les langues Mon-Khmer. *The Mon Khmer Studies Journal* 8:1-76.
- Gafos, Adamantios & Benus, Stefan. 2006. Dynamics of phonological cognition. *Cognitive Science* 30(5), 1-39.
- Gregerson, Ken and Smith K. 1973. The development of Tōdrah register. *The Mon-Khmer Studies Journal* 4:143-184.
- Haudricourt, André. 1965. Les mutations consonantiques de occlusives initiales en Mon-Khmer. *Bulletin de la Société Linguistique de Paris* 60:160-172.
- Henderson, Eugénie. 1952. The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies* 14(1):149-174.

- Huang, Bufan. 1995. Conditions for tonogenesis and tone split in Tibetan dialects. *Linguistics of the Tibeto-Burman Area* 18:43–62.
- Huffman, Franklin. 1976. The register problem in fifteen Mon-Khmer languages. *Oceanic Linguistics Special Publications* 13:575-589.
- Hyman, Larry. 1976. Phonologization. *Linguistic studies presented to Joseph H. Greenberg*, ed. by Juilland Alphonse, 407-418. Anma Libri.
- Johnson, Keith. 1997. *Acoustic and Auditory Phonetics*. Blackwell.
- Keating, Patricia. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60(2):286-319.
- Kent, Raymond, and Read, Charles. 2002. *Acoustic Analysis of Speech*. Singular-Thomson Learning.
- Kingston, John, and Randy Diehl. 1994. Phonetic knowledge. *Language* 70:419-454.
- Kingston, John, and Neil Macmillan. 1995. Integrality of nasalization and F1 in vowels in isolation and before oral and nasal consonants: A detection-theoretic application of the Garner paradigm. *Journal of the Acoustical Society of America* 97:1261-1285.
- Lisker, Leigh. 1986. "Voicing" in English: Catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29:3-11.
- Maclagan, Margaret, and Jennifer Hay. 2007. Getting fed up with our feet: Contrast maintenance and the New Zealand "short" front vowel shift. *Language Variation and Change* 19:1-25.
- Matisoff, James. 1973. Tonogenesis in Southeast Asia. *Consonant Types and Tone: Southern California Occasional Papers in Linguistics 1*, ed. by Larry Hyman, 71-96.
- Ohala, John. 1993. The phonetics of sound change. *Historical Linguistics: Problems and Perspectives*, ed. by Charles Jones, 237-278. Longman.
- Phaen, T., Sok, M., & Bequette, T. 2012. *Bunong-Khmer Bilingual Dictionary*. Mondulhiri Province, Cambodia: International Cooperation Cambodia, READ Project.
- Phillips, Richard. 1973. Mnong vowel variations with initial stops. *The Mon-Khmer Studies Journal* 4:119-127.
- Repp, Bruno. 1986. Comment. *Invariance and Variability in Speech Processes*, 449-455. Lawrence Erlbaum Associates.
- Stevens, Kenneth, Samuel Keyser, and Haruko Kawasaki. 1986. Toward a phonetic and phonological theory of redundant features. *Invariance and Variability in Speech*

Processes, ed. by Perkell, Joseph, and Dennis Klatt, 426-449. Lawrence Erlbaum Associates.

Turk, Alice, and James Sawusch. 1996. The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America* 99:3782-3790.

Vogel, Sylvain and Jean Michel Filippi. 2006. *Éléments de langue Phnong*. Editions Funan.

Wayland, Ratre and Allard Jongman. 2003. Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics* 31:181-201.

Whalen, D. H., and Andrea Levitt. 1995. The universality of intrinsic f0 of vowels. *Journal of Phonetics* 23:349-366.

Appendix 1: Wordlist for Experiment 1

		1	2	3	4	5	6
		p	b	t	d	k	g
1	i	pih	bih	tih	dih	king ngang	gi
2	ɪ	pě	bě	tĩk	dĩk	kĩt	gĩt
3	e	pêh	bên	têk	dadê	kêng mang	gê
4	ɛ	pe	be	te jung	děh	kět	gěh
5	ɨ	pɨ	bɨt	tɨm	dɨm	kɨ	gɨt
6	ə	pəh	bəh saəm	təm	dəm	kət	gəm
7	u	pu	būt	tu	du	kuh	gū
8	o	poh	bo	toch	dol	kor	go
9	ɔ	pɔ	bɔh	tɔh	dɔh	kɔp	gɔp
10	a	pal	bal	tak	dak	kan	gak

Table 7: Production wordlist

Appendix 2: Results of Experiment 2

token #	onset	VOT	vowel	pitch	Participant Choices and Goodness Ratings									
					1		2		3		4		Avg	
1	R2	R2	R1	R2	R1	2	R1	3	R1	7	R1	7	R1	4.8
2	R2	R2	R1	R1	R2	3	R1	4	R1	7	R1	7	R1	5.3
3	R2	R1	R1	R2	R1	2	R1	5	R1	7	R1	7	R1	5.3
4	R2	R1	R1	R1	R2	3	R1	7	R1	7	R1	7	R1	6
5	R1	R2	R2	R2	R1	7	R2	5	R2	7	R2	7	R2	6.5
6	R1	R2	R2	R1	R2	3	R2	6	R2	7	R2	6	R2	5.5
7	R1	R1	R2	R2	R1	7	R2	7	R2	7	R2	7	R2	7
8	R1	R1	R2	R1	R1	7	R1	6	R2	7	R2	5	R2	6.3

Table 2: Exp 2 Responses- High Vowels

token #	onset	VOT	vowel	pitch	Participant Choices and Goodness Ratings									
					1		2		3		4		Avg	
9	R2	R2	R1	R2	R1	2	R1	6	R1	4	R1	5	R1	4.3
10	R2	R2	R1	R1	R1	2	R2	7	R1	4	R1	4	R1	4.3
11	R2	R1	R1	R2	R1	3	R2	6	R1	4	R1	4	R1	4.3
12	R2	R1	R1	R1	R1	7	R2	7	R1	4	R1	5	R1	5.8
13	R1	R2	R2	R2	R2	7	R2	5	R2	4	R2	7	R2	6.5
14	R1	R2	R2	R1	R2	3	R2	7	R2	7	R2	7	R2	6
15	R1	R1	R2	R2	R2	7	R1	7	R2	7	R2	7	R2	7
16	R1	R1	R2	R1	R2	3	R1	6	R2	7	R2	7	R2	5.8

Table 3: Exp 2 Responses- Mid Vowels

token #	onset	VOT	vowel	pitch	Participant Choices and Goodness Ratings									
					1		2		3		4		Avg	
17	R2	R2	R1	hi	R1	2	R2	3	R1	7	R1	7	R1	4.8
18	R2	R2	R1	low	R1	7	R2	4	R1	7	R1	7	R1	6.3
19	R2	R2	R1	R1	R1	4	R1	2	R1	7	R1	7	R1	5
20	R2	R1	R1	hi	R1	2	R1	7	R1	7	R1	7	R1	5.8
21	R2	R1	R1	low	R1	4	R1	6	R1	7	R1	7	R1	6
22	R2	R1	R1	R1	R1	3	R1	4	R1	7	R1	7	R1	5.3
23	R1	R2	R2	hi	R2	5	R2	7	R2	7	R2	7	R2	6.5
24	R1	R2	R2	low	R2	2	R2	6	R2	7	R2	7	R2	5.5
25	R1	R2	R2	R2	R2	7	R2	6	R2	7	R2	7	R2	6.8
26	R1	R1	R2	hi	R2	5	R2	6	R2	7	R2	7	R2	6.3
27	R1	R1	R2	low	R2	2	R2	6	R2	7	R2	7	R2	5.5
28	R1	R1	R2	R2	R2	7	R1	5	R2	7	R2	7	R2	6.5

Table 4: Exp 2 Responses- Low Vowels