TONOGENESIS IN KHMER

Ratree Wayland
University of Florida
<ratree@ufl.edu>

Susan Guion
University of Oregon
<guion@uoregon.edu>

0 Introduction

Physiological constraints of the articulatory and/or auditory mechanisms have been proposed as the source of sound changes for some time (e.g., Hombert 1977 and references therein, Ohala 1971, 1974, 1981b, 1989, 1993). As pointed out by Hombert (1977), these explanations imply that the speaker's pronunciation may not be perceived as intended. The distortion may occur due to articulatory and auditory constraints, which affect the way the sounds are produced and perceived by listeners. Ohala (1993) reviews many sound changes and proposes a typology of sound change in which variation in speech due to coarticulation (among other things) can fail to be corrected by the listener (hypo-correction) or corrections can be erroneously applied (hyper-correction). As coarticulations or reductions are greater in faster speech (e.g., Guion 1998, Moon, Lindblom & Lame 1995, Lindblom 1990), one might expect more hypo-correction in fast speech forms. Preliminary evidence for such a proposal was given in Guion (1998) in which faster speech forms were found to be more similar acoustically to post sound change forms than citation speech forms.

The goal of this paper is to report on tone development in Khmer (Cambodian). Based on previous research, it will be proposed that this development is largely phonetically motivated. Moreover, we propose that the sound change has its origins in faster, colloquial speech.

1 Tone Development in Khmer

Recently, Thach (1999) has reported a sound change in dialects of Khmer spoken in Vietnam in which consonant + [r] clusters in onset position of main syllables lose the [r] and gain a falling tone on the following vowel (e.g., [krɒː] > [kɒ̀ː] 'poor'). The sound change is quite advanced in these dialects, especially among younger speakers. Through this sound change, tone has been introduced and there are now minimal pairs such as [kɒ̀ː] 'poor' (from [krɒː]) and [kɒː] 'neck'. Additionally, Thach (1999) reports that the trill [r] has become a glottal fricative [h] in syllable initial position in main syllables. Informal observation of the Khmer spoken in Cambodia suggests that a similar sound change is happening there as well.

The sound change in Khmer spoken in Cambodian to be investigated here involves monosyllabic words[1] with a consonant cluster onset. The first consonant in the cluster is an obstruent and the second member is an alveolar trill [r]. This sound change occurs only in colloquial pronunciation, but not careful reading pronunciation. In colloquial pronunciation, the alveolar trill [r] in a voiceless obstruent + [r] sequence is dropped or becomes a fricativized voiceless aspirated [r]. In most cases, the initial voiceless unaspirated stop also changes into a voiceless aspirated stop. If the vowel involved is a

Ratree Wayland, John Hartmann & Paul Sidwell, eds. *SEALSXII: papers from the 12th meeting of the Southeast Asian Linguistics Society (2002)*. Canberra, Pacific Linguistics, 2007, pp.145-152.

short low front [a] vowel, it becomes the diphthong [ea]. More importantly, a distinct rising pitch contour is also evident in words having undergone the changes.

2 Acoustic Analysis

The goal of this analysis is to determine acoustically whether or not the above mentioned sound changes have indeed occurred. To this end, an acoustic analysis of a subset of words undergoing these changes is undertaken.

2.1 *Method*

Speaker: A male native speaker of Khmer from Phnom Penh served as the speaker. He is in his 40's and has been living in the United States for approximately 5 years at the time of recording. No known speech or hearing impairment was reported.

Stimuli: Stimuli were 13 Khmer words (see Appendix) elicited from the speaker. Nine out of the thirteen words begin with a voiceless stop consonant, three with a voiceless affricate and one with a voiceless fricative. The speaker was instructed to read the words from the wordlist at a normal speaking rate. For comparison purposes, the speaker was instructed to first say the word as he normally would if he were to read it from a book (spelling or reading pronunciation), and second as he would in a conversation to another Khmer speaker (colloquial pronunciation). Each word was repeated three times in random order.

2.2 *Procedure*

Recording of the wordlist was digitized on a Kay Elemetrics CSL station (Model 4300B) at a sampling rate of 25 kHz. Each word was edited and stored as a separate file for further analysis using Cool-Edit (www.syntrillium.com). Subsequent acoustic analyses were performed using Kay Elemetrics MultiSpeech. Acoustic measurements included voice-onset-time (VOT) for words with stop initials, frication duration for words with affricate initials, vowel fundamental frequency (F_0) and frequencies of the first and second formants (F1, F2).

2.3 *Results*

VOT: Mean VOT duration of initial stop consonants for both modes of pronunciation is shown in Table 1. With the exception of [krɒː] 'poor', VOT duration of initial stops in colloquial speech was generally longer than that of the reading pronunciation (mean = 44 vs. 27 ms). Result of a paired t-test supported this observation ($t = 2.06$, $p < .0001$, $df = 26$). This finding confirmed that, in most cases, a voiceless stop [p, t, k] initials in reading pronunciation became aspirated in colloquial pronunciation.

Table 1: *Mean VOT duration (in ms) for stop initials in reading and colloquial modes of pronunciation.*

Word	Colloquial	Reading
[pram] ‘five’	39	13
[praə] ‘use’	31	15
[prap] ‘tell’	34	13
[pruəy] ‘sad’	42	33
[priəy] ‘spirit’	34	21
[trɨw] ‘correct’	52	45
[kru] ‘teacher’	51	36
[krɒː] ‘poor’	50	51
[krŏən] ‘enough’	66	21
Mean	**44**	**27**

Fricative Noise Duration: Mean fricative noise duration after the burst of a voiceless affricate initial is shown in Table 2. As expected, for all three words, fricative noise duration was longer for colloquial than for reading pronunciation (mean = 82 vs. 66 ms.). This difference was significant in a paired t-test (t = 2.31, p <.01, df = 8). This finding suggested that a voiceless affricate [tç̱] in reading pronunciation became a voiceless aspirated affricate [tç̱^{h}] in colloquial speech.

Table 2: *Mean fricative duration (in ms) of voiceless affricate [tç̱] in reading and colloquial pronunciation.*

Word	Colloquial	Reading
[tç̱roʊt] ‘harvest’	82	69
[tç̱rɨw] ‘deep’	80	62
[tç̱rŏəm] ‘muddy’	84	67
Mean	**82**	**66**

As for [s] in [srəy] ‘female’, an auditory as well as a spectrographic examination (Figure 1 and 2) suggested that, in most cases, it became a voiceless aspirated affricate [tç̱^{h}].

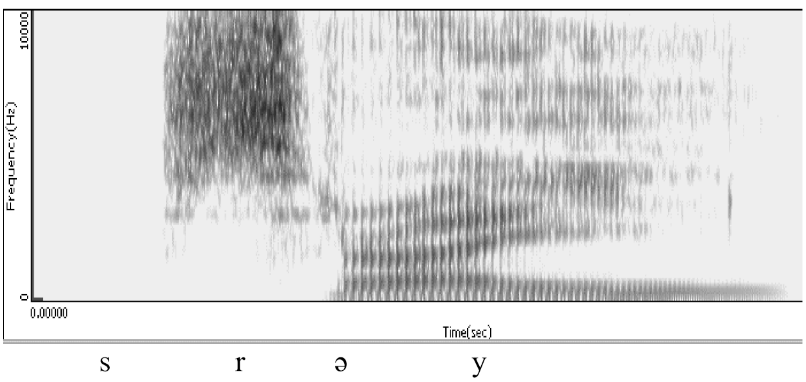

Figure 1: *Spectrogram of [srəy] ‘female’ in reading pronunciation.*

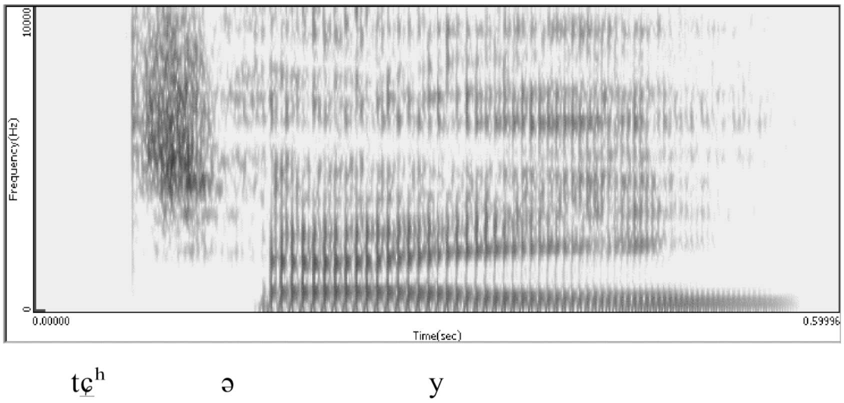

Figure 2: *Spectrogram of [srəy] 'female' in colloquial pronunciation.*

It should be mentioned, however, that the replacement of the alveolar trill [r] with aspiration was not always complete. In some cases, it became a fricativized voiceless aspirated [r].

F_0 Contour: F_0 contours (in Mels) for colloquial and reading pronunciation averaged across all words is illustrated in Figure 3. In general, F_0 values at vowel onset (182 vs. 181 mels) and at vowel offset (244 vs. 249 mels) are comparable in both types of pronunciation. However, while F_0 continues to rise in reading pronunciation, there is a sharp decrease in F_0 value at 50% in the vowel for colloquial speech pronunciation (t = 2.18, p <.006, df =12). Thus, while rising pitch contour is the characteristics of reading pronunciation, colloquial pronunciation is best characterized with a falling-rising contour.

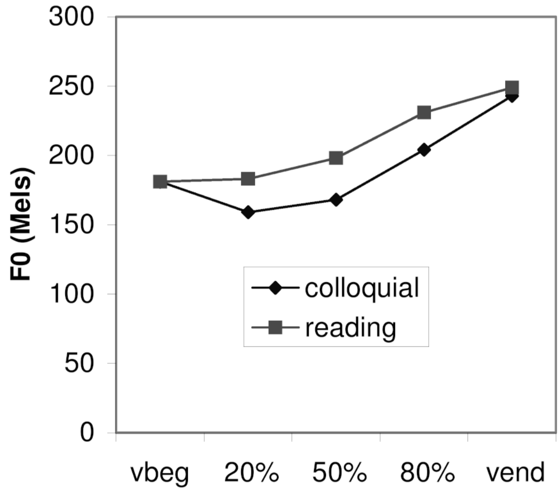

Figure 3: *Normalized F_0 contours for reading and colloquial pronunciations.*

A Vowel Quality Change: The change from [a] to [ea] is evident in the spectrograms derived from the two different pronunciations of [prap] 'to tell' included in Figure 4 and 5. The locations and movements of the first two formants of the vowels in the two spectrograms clearly showed a change from a monophthong [a] to a diphthong [ea].

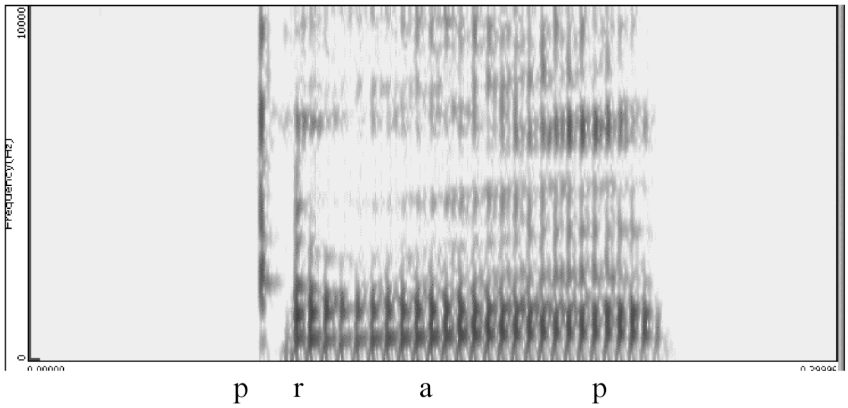

Figure 4: *Spectrogram of [prap] 'tell' in reading pronunciation.*

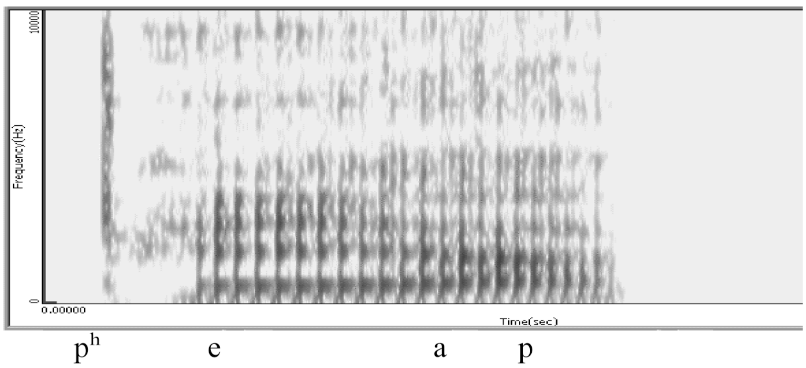

Figure 5: *Spectrogram of [prap] 'tell' in colloquial pronunciation.*

This observation is supported by acoustic measurements of the formant frequencies of the two vowels in all three tokens of [pram] 'five' and [prap] 'tell'. The first formant (F1) and second formant (F2) values of the vowel in [pram] 'five' in the reading pronunciation is 832 Hz and 1364 Hz, and 766 Hz and 1248 Hz for [prap] 'tell' respectively.

In colloquial pronunciation, mean F1 and F2 values of the first element of the diphthong in [pram] 'five' are 572 Hz. and 1816 Hz, while those of the second element are 861 Hz. and 1382 Hz. Those of [prap] 'tell' are 504 Hz, and 1819 Hz vs. 706 Hz and 1586 Hz. respectively. Thus, it is evident that the [pram] 'five' and [prap] 'tell' in reading pronunciation were pronounced as [p^heam] and [p^heap] in colloquial speech.

In summary, the following sound changes occurred when words beginning with an obstruent + r sequence are produced in colloquial style. (1) The alveolar trill [r] was either dropped or became a fricativized voiceless aspirated [r]. (2) In most cases, voiceless stop [p, t, k] becomes voiceless aspirated stops [p^h, t^h, k^h]. (3) Voiceless affricate [tɕ] and voiceless fricative [s] become voiceless aspirated affricate [tɕ^h]. (4) Low vowel [a] becomes the diphthong [ea]. (5) The whole syllable is now produced with a falling-rising pitch contour. Note that this finding is somewhat different from the Khmer dialect spoken in Vietnam reported in Thach (1999). Perhaps the falling tone reported for this dialect is a description of the falling part of the contour, or perhaps the tone development is different in the two dialects of Khmer. An acoustic investigation of the Khmer dialect spoken in Vietnam would need to be undertaken to decide the matter.

In the next section, we turn to a discussion on plausible phonetic explanations for these sound changes in Khmer.

3 A Phonetic Account

We hypothesize that all the changes are induced by a relatively faster speaking rate in colloquial speech in comparison to a more careful reading pronunciation. First, we propose that the alveolar trill [r] becomes a voiceless [r̥] and assimilates its voicing feature with the preceding voiceless obstruents. This may be due to the temporal overlap or carry-over of the glottal gestures of the initial voiceless obstruent into the following [r]. That is, the vocal folds remain open throughout the initial cluster and do not adduct to begin voicing until the onset of the following vowel. It is also possible that, for aerodynamic reasons, voicing during a trill [r] is difficult to achieve. A relatively strong and high volume of airflow is needed to set the tongue tip into the trill motion. However, this can only be achieved, especially in a short period of time, if the glottis is relatively wide open. In turn, the wide-open glottis makes voicing difficult. Either of these potential explanations could have a plausible origin in faster, colloquial speech if a view of speech production is taken in which energy expenditure on the speaker's part tends toward economy to the extent allowed by the speaking conditions (see e.g., the Hypo and Hyper-Speech theory of Lindblom 1990). On this view, the costly processes of laryngeal timing or aerodynamic adjustments are not reliably used in casual speech. Presumably, in casual speech, there is more signal redundant information allowing the lower cost form of behavior while still permitting sufficient information transfer for effective communication.

In the next step in the sound change, the voicelessness of the [r] is heard as aspiration and becomes perceptually associated (phonologized) with the initial obstruent consonants. This process accounts for the change from voiceless unaspirated stops and affricates to voiceless unaspirated stops and affricates. When producing aspiration[2], it is advantageous for the size of the oral constriction aperture of the initial consonants to be decreased. The smaller or complete closure of the oral cavity will result in an increase in oral pressure and thus, a delay in obtaining a transglottal pressure drop suitable for voicing. This mechanism results in a change from an alveolar fricative [s] to a voiceless aspirated affricate [tɕ^h].

Another aerodynamic measure adopted by Khmer to insure a desired degree of aspiration is the tongue raising gesture at the onset of the vowel, leaving only a small oral constriction aperture, thereby delaying the VOT. This explains why the low short vowel [a] becomes a diphthong [ea], as in [p^{h}eam] 'five' and [p^{h}eap] 'tell'. This mechanism appears to be needed only when a low vowel is involved. A strong degree of aspiration achieved through these aerodynamic mechanisms may also be responsible for a slight degree of perceived breathiness in these syllables

A change in F_0 pattern in colloquial pronunciation may also be attributed to the aerodynamic factor. As shown in Figure 3, there is a significant drop in F_0 value from the beginning of the vowel to vowel mid-point resulting in a falling-rising pitch contour for syllables produced with this style of pronunciation. The decrease in F_0 may be due to the cessation or reduced rate of airflow when voicing begins after aspiration. That is, the reduction of high volume airflow results in a weak Bernoulli effect, thus a slower rate of vocal fold vibration. This effect was not found in reading pronunciation where initial obstruents are not aspirated.

In conclusion, all sound changes in Khmer including the aspiration of initial obstruent consonants, a change from a voiceless fricative [s] to a voiceless aspirated affricate [tɕʰ], and a change in vowel quality from [a] to [ea] may be attributed to an aerodynamic effect. All of these changes occur as a result of aerodynamic mechanisms adopted to secure the desired degree of aspiration in colloquial pronunciation once the aspiration, which has its origin from the voiceless trill [r̥] is phonologized and becomes part of the initial consonant. A reduction in high volume airflow after aspiration cause the F0 to drop once voicing begins leading to a distinct falling-rising pitch contour.

Notes

The authors would like to thank Sopheak Son for the stimuli used in the study. Many thanks also go to Mohamed Al-Khairy for obtaining some of the references cited.

1. Even though the changes also occurred in second syllables of bi-syllabic words such as [baːrəy] 'cigarette', [bɒŋriən] 'teach' etc., they will not be included here. Bi- syllabic words often become sesqui-syllabic words e.g., [bərəy] or [pərəy] 'cigarette' or [pəriən] 'teach' in colloquial pronunciation. The main (second) syllable which begins with an alveolar trill [r] will undergo the same changes to be described here. The phonetic explanation proposed here should be applicable to those cases as well.

2. There is evidence that the aspiration is further strengthened and becomes [h] in bi-syllabic words. For example, [bɒŋriən] 'to teach' is also heard as [pəhiən] (Pisitpanporn, 1999), or in monosyllabic words as in [riən] > [hiən] 'to learn' (Thach, 1999).

References

Guion, S. G. (1998). The role of perception in the sound change of velar palatalization. Phonetica 55, 18-52.

Hombert, J. M. (1977). Development of tones from vowel height? *Journal of Phonetics* 5, 9-16.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory in Mardcastly, W. & Marchal, A., (eds.), *Speech production and speech modeling*, 403-439. Dordrecht: Kluwer.

Moon, J-S., Lindblom, B., & Lame, J. (1995). A perceptual study of reduced vowels in clear and casual speech. In Elenius, K., & Branderud, P., (eds.), 670-677. Proceedings ICPhS 95, vol 2., Stockholm.

Ohala, J. J. & Ewan, W. G. (1973). Speed of pitch change. *Journal of the Acoustical Society of America* 53, 345.

Ohala, J. J. (1974). Phonetic explanation in phonology. In Bruck, Fox, LaGaly (eds.), *Papers from the parasession on natural phonology*, 251-74, Chicago: Chicago Linguistic Society.

_____. (1971). The role of physiological and acoustic models in explaining the direction of sound change. *Project on linguistic analysis: Reports, Second Series*, no 15.

Phonology Laboratory—Department of Linguistics, University of California, Berkeley

_____. (1975). Conditions for vowel devoicing and frication. *Journal of the Acoustical Society of America* 58, S39.

_____. (1976). A model of speech aerodynamics. *Report of the Phonology Laboratory* (Berkeley) 1, 93-107.

_____. (1978). Production of tone. In Fromkin, V. (ed.), *Tone*, 5-32. United Kingdom: Academic Press.

_____. (1981a). Articulatory constraints on the cognitive representation of speech. In Myer, T., Laver, J. and Anderson, J. (eds.), 111-122. *The cognitive representation of speech*. North-Holland Publishing Company: Amsterdam, New York, Oxford.

_____. (1981b). The listener as the source of sound change. In Masek, Hendrick, Miller, (eds.), 178-203. *Papers from the parasession on language and behavior*, Chicago Linguistic Society: Chicago.

_____. (1989). Sound Change is drawn from a pool of synchronic variation. In Breivik, Jahr, (ed.), 171-198. *Language change: contributions to the study of its causes*, NY: Mouton de Gruyter.

_____. (1993). The phonetics of sound change; in Jones, D. (ed.), *Historical linguistics: problems and perspectives*, 237-278, London/NY: Longman.

Thach, N. M. (1999). Monosyllabization in Kiengiang Khmer. *Mon-Khmer Studies* 29, 81-95.

Appendix: Wordlist

[pram]	“five”
[praə]	“to use”
[prap]	“to tell”
[pruəy]	“sad”
[priəy]	“spirit”
[trɨw]	“correct”
[kru]	“teacher”
[krɒː]	“poor”
[krŏən]	“enough”
[srəy]	“female”
[tɕ̠roʊt]	“to harvest (rice)”
[tɕ̠rɨw]	“deep”
[tɕ̠rŏəm]	“muddy”