

## Sound change in Khmer: Experimental and computational studies

**Introduction** Unlike many languages of Southeast Asia, Khmer (Cambodian) is not a tone language. However, linguists have noted a pitch-based contrast between certain words in the colloquial speech of a number of Khmer dialects since at least the 1960s (e.g. Noss, 1966). While the emergence of lexical tone is common in languages of Southeast Asia, the manner by which it might be taking place in Khmer – as a result of loss of /r/ in initial clusters – has not been reported for any other language. This paper presents new acoustic and perceptual data on the emergence of F0-based contrast in Phnom Penh Khmer, and shows how a computational perspective can help us understand the dynamics of this type of sound change.

**Background** In Standard Khmer, one can find minimal triplets such as /kɑɑ/ ‘neck’, /k<sup>h</sup>ɑɑ/ ‘donkey’ and /kraɑ/ ‘poor’. In the colloquial speech of the capital Phnom Penh, /r/ is lost in /Cr/ onset clusters, but several sources (e.g. Noss, 1966; Wayland & Guion, 2005) report that forms like ‘poor’ are instead distinguished by other acoustic cues such as aspiration, a falling-rising pitch contour, breathy voice quality, and in some cases diphthongization, e.g. /kraɑ/ > [k<sup>h</sup>ǎɑ] but /kruu/ ‘teacher’ > [kũu]. However, the only acoustic description of this phenomenon (Wayland & Guion, 2005) is preliminary, and it is not clear which of these cues have become perceptually sufficient or salient for listeners.

**Experiments** 20 native speakers of Phnom Penh Khmer participated in one production and two perception tasks. In the production task, subjects read a wordlist of minimal /CV(C), C<sup>h</sup>V(C), CrV(C)/ triplets covering a range of vowel qualities in both careful and casual conditions. Subjects also participated in two 2AFC listening experiments, designed to test the perceptual salience of the acoustic cues present in the colloquial forms. For the first task, a 7-step [kuu ~ kũu] continuum varying in F0 midpoint was synthesized using a Klatt synthesizer, and used as the basis for two additional continua by adding a fixed degree of aspiration [k<sup>h</sup>uu ~ k<sup>h</sup>ũu] or breathy voice [kũu ~ kũu]. For the second task, a 7-step [kɑɑ ~ koɑ] continuum varying in F1 height was used as the basis for three additional continua, created by adding fixed aspiration [k<sup>h</sup>ɑɑ ~ k<sup>h</sup>oɑ], breathy voice [kɑɑ ~ koɑ], or a falling-rising F0 contour [kãɑ ~ kõɑ]. This design allowed the potentially additive effect of a variety of acoustic cues to be explored with a minimum number of trials.

**Results** Perceptual response data suggest that F0 has become the primary cue used by listeners to identify colloquial /CrV(C)/ forms, although other cues (notably aspiration) also play a role. Analysis of the production data shows colloquial productions of /CrV(C)/ forms are characterized by a falling-rising F0 contour and increased post-release aspiration, with a shift in F1 limited to forms containing low vowels. Crucially, careful productions of the same forms are found to contain excrescent vowels, *contra* Huffman (1972). I will suggest that the F0 contour finds its source in the higher pitch associated with the excrescent vowel, while the restricted distribution of diphthongization results from the perceptual interaction of aspiration (perceived as breathiness) with vowel height. This result is predicted from a model-based clustering approach to phonological categories (Kirby, 2010, 2011), whereby bias (here, devoicing of /r/ in onset clusters) led to probabilistic enhancement of F0.