**Illusory vowels and illusory tones in the perception of consonant clusters by monolingual Chinese Mandarin speakers**

Qianwen Guan (qianwen.guan@linguist.univ-paris-diderot.fr)
CLLILAC-ARP, Université Paris Diderot, 75013 Paris

Previous work shows that listeners tend to perceive an illusory vowel inside consonant clusters that are illegal in their native language [1],[2]. But few studies have been concerned with the perception of tones in connection with L2 phonotactics, specifically for L1 speakers of tone languages. This study examines how L1 speakers of a tone language (Mandarin) perceive the clusters of an L2 language without tone (Russian). The issue that we address is how the perception of L2 phonotactics relates to the perception of tones. Some studies of loanwords showed that speakers of languages preferring simple CV syllable structure insert an illusory vowel as well as an illusory tone inside illegal consonant clusters [5], [6].

Our study investigates experimentally the interaction of phonotactic and prosodic adaptation for Mandarin L1. For this purpose we tested the perception of non-native clusters by monolingual Mandarin speakers. The experiment was conducted in Beijing, with 24 participants especially recruited from among monolingual speakers with minimal exposure to a foreign language. The stimuli were non-words produced by a Russian native speaker, containing CC sequences and controls (see table 1). The participants were asked to transcribe each stimulus they heard in Pinyin, including tones. Different types of transcription errors (e.g. epenthesis, metathesis) were observed and analyzed statistically.

In this abstract I report only the general result, and I focus on the transcription of tones. The most frequent transcription included vowel epenthesis in the clusters. The illusory vowel is always perceived as short central [ɤ] (*akta->akɤ̲ta*), except after a labial (*ipta->ipu̲ta*). In the controls, the vowel /a/ between clusters is perceived as /a/ when it precedes a stressed syllable (*ka̲tápa*). When it follows a stressed syllable (*áka̲ta*), it is perceived as central /ɤ/. This response reflects the fact that in Russian, an unstressed vowel preceding stress is less reduced than an unstressed vowel elsewhere in the word [3]. Thus, Mandarin listeners can still perceive a moderately reduced [a] in pre-tonic position, but report a fully reduced vowel in post-tonic position. Of the 24 participants, only 8 could transcribe tones. This is because Pinyin transcription, though commonly used, rarely includes tones. Moreover, the participants do not have more than high school education and do not write regularly. The results of the 8 speakers are summarized and discussed below.

*Word position and position with respect to stress.* In the Russian stimuli, the stressed syllable is realized with an F0 peak. The listeners' responses thus correspond to the low F0 in the pre-stressed syllable, and to the falling F0 on the post-stressed syllable, respectively (fig. 1).

*Vowel quality*. Overall, vowel quality, whether of illusory vowels or of real vowels, does not influence the perception of illusory tones ($p>0.05$).

*Consonant type*. There is a significant difference of tone perception in relation to phonotactics in initial, pre-stressed clusters. Listeners reported high tone on the illusory vowel significantly more often in stop-stop *(e.g., **kt**ápa)* than in stop-nasal clusters ($p<0.05$). While the difference is not significant between stop-stop and stop-liquid clusters ($p= 0.067$), the observed tendency is still for more high tone in stop-stop clusters (fig. 2). This difference is not found in the controls ($p>0.05$), on a real vowel. In the clusters, where no vowel is actually produced, the audible F0 movement is very fast, rising toward the peak of the stressed vowel. The start of the F0 rise is influenced by the consonant type [4]: after a voiceless stop F0 rises in the following vowel, while after a sonorant stop F0 falls. In the controls *(ka̲tápa),* F0 starts low on the unstressed vowel, and rises into the following stressed syllable. The participants seem to respond to this systematic rise, rather than to the subtler effects of consonant-type.

The results suggest that the perception of illusory tones as prosodic adaptation in L2 stimuli accompanies the perception of illusory vowels as phonotactic adaptation. In both phonotactic and prosodic adaptation, listeners show sensitivity to fine acoustic phonetic detail.

**Annexes:**

| | **Word-initial (pre-stressed)** | | **Word- medial (post-stressed)** | |
|---|---|---|---|---|
| CV structure | *#CVCVCV* | katápa, takápa, patáka, paláka, kalápa, kanápa | *VCVCV* | ákata, átaka, ípata, átapa, ákanu, ákalu, ípala |
| CC cluster | *#CCVCV* | ktápa, tkápa, ptáka, pláka, klápa, knápa | *VCCV* | ákta, átka, ípta, átpa, áknu, áklu, ípla |

**Table 1**: The stimuli contain two structures (CV structures and CC clusters) and two positions (word-initial and word-medial). Stress falls after the cluster when word- initial *(#katápa, #ktápa*), and before the cluster when word-medial (*ákata, ákta*). In the cluster C1 is always a voiceless stop and C2 is either a voiceless stop ([p,t,k], [n], or [l]. The vowel preceding a cluster is /a/ or /i/ (e.g. ***a***kta, *iknu*). The vowel in CV structures (*k**a**tapa*) is always /a/.

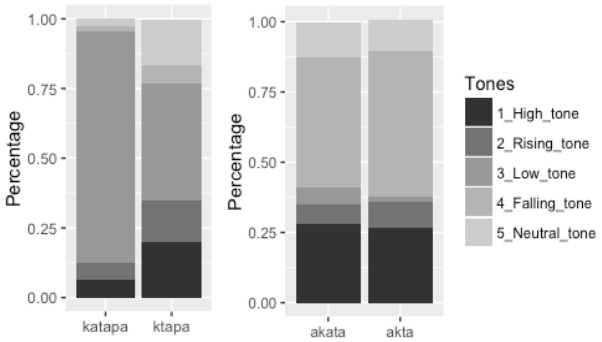


**Figure 1**: Percentage of perceived tones in initial (pre-stressed) syllable (*katápa, ktápa*) and in medial (post-stressed) syllable (*ákata, ákta*). Low tone was mostly perceived in an initial, pre-stressed syllable (*katápa, ktápa*) and falling tone in a medial, post-stressed syllable (*ákata, ákta*) on both real and illusory vowels ($p$<0.001).

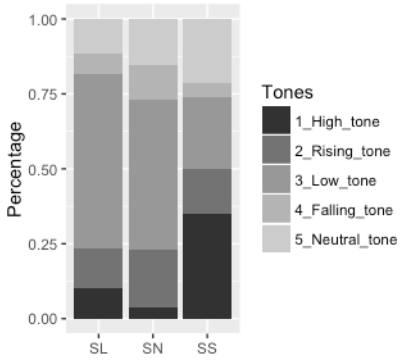


**Figure 2**: Percentage of perceived tones in the initial (pre-stressed) cluster with three consonant types: stop+liquid SL, stop+nasal SN, stop+stop SS. Listeners perceived more high tone significantly in SS clusters than in SN clusters ($p$<0.05). While there is no significant difference between SS clusters and SL clusters (p= 0.067), the observed tendency is still for more high tone in SS clusters.

**References:**

[1] Berent, I., Lennertz, T., Jun, J., Moreno, M. A., Smolensky, P. (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences*, 105(14), 5321–5325. [2] Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25 (6), 1568- 1578. [3] Hamilton, W.S (1980) Introduction to Russian phonology and word structure. Columbus: Slavica
[4] Hombert, Jean-Marie (1978). Consonant types, vowel quality, and tone. In Fromkin (1978). 77–111. [5] Silverman, D. (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 9. 289 – 328. [6] Ufomata, T. (1991). Englishization of Yoruba phonology. *World English* 10, 33-51