Native listeners perceive illusory vowels when presented with sound sequences that do not respect the phonotactic constraints of their language [1,2]. More specifically, when a native listener is presented with a word-medial consonant sequence that violates the phonotactic constraints in their language, an illusory vowel is perceptually induced in between such sequences thereby creating an illusory sequence that respects the phonotactic constraints of the language. For e.g., when a Japanese listener is presented with [ebzo], they may actually perceive [ebuzo] given that [bz] is an illicit consonant sequence in Japanese. With respect to the quality of the illusory vowel, it has been argued that it can only be the phonetically minimal vowel in the language [3], and that multiple illusory vowels cannot be perceived during speech perception [4]. Here, we show that there can indeed be different illusory vowels in different contexts, and that the quality of the illusory vowels is itself modulated by the phonological patterns in the language.

Mandarin has a phonotactic constraint preventing non-nasal coda consonants, so sequences of obstruent consonants followed by nasal consonants are disallowed in the language [*teçm, *tçm]; this is true both within and across words. Furthermore, alveopalatal consonants [tç, tçh, ç] can only appear before high front vowels or glides [i, j, y, ü], while alveolar stops can precede any vowel. Therefore, when Mandarin listeners are presented with a phonotactically illicit sequence involving an alveopalatal consonant (e.g., [tçh]), we predict that the only vowels that can be inferred to maintain the phonotactic patterns of the language are high front vowels. However, when they are presented with a phonotactically illicit sequence involving an alveolar stop (e.g., [tçm]), there is no phonotactic bias for any particular vowel; in this context, we expect the phonetically minimal (default) vowel [a] to be the illusory vowel. Contrastingly, for American English listeners, [tç] is expected to be perceived as [ʃ], and both [tçm] and [tçm] are phonotactically valid sequences (especially, within compounds and across words); therefore, the phonological patterns in the language do not promote the perception of illusory vowels for English listeners; however, there could be phonetic factors in the stimuli themselves that could induce illusory vowels even for English listeners. We present two experiments on Mandarin listeners with American English listeners acting as controls (for phonetic artifacts in the stimuli) in support of the above predictions that Mandarin listeners perceive different illusory vowels in different contexts.

**Experiment 1:** We ran an ABX experiment on 20 native Mandarin listeners, and 19 native American English listeners as controls. We presented participants with pairs of nonce words of the form aC1V1ma [C1 = tç / tçh; V1 = i / a / Ø (Null)]. All tokens had initial stress, and recorded by a trained phonetician. We take confusability between pairs of words with and without vowels to suggest the induction of an illusory vowel. A mixed ANOVA of the percentages of correct responses revealed a main effect of language [F(1, 37) = 8.37, p <.001], a main effect of word pair [F(5, 185) = 10.63, p <.001], and an interaction of word pair by language [F(5, 185) = 15.36, p <.001]. We ran planned t-tests to compare the percentage of correct responses for each relevant word-pair between the two language groups (Figure 1).

![Figure 1: Exp. 1 - Average % correct for English and Mandarin listeners (error bars = 1 S.E.; ** = p<0.01)](image-url)
Results suggest that, as expected, Mandarin listeners were confusing not only \([\text{at}^b\text{ima-}\text{at}^b\text{ma}]\) and \([\text{ate}^b\text{ima-}\text{ate}^b\text{ma}]\) more than English listeners. But, they were also confusing \([\text{ate}^b\text{oma-}\text{ate}^b\text{ma}]\) more than English listeners unexpectedly. Furthermore, English listeners performed worse with \([\text{at}^b\text{oma-}\text{at}^b\text{ima}]\); potentially due to the effect of the phonological process of vowel reduction in unstressed positions.

**Experiment 2:** The unexpected confusions observed for the Mandarin listeners for the crucial word pairs in Exp. 1 could have been due to other perceptual changes. For e.g., \([\text{ate}^b\text{oma}]\) might have been perceived as \([\text{ate}^b\text{ima}]\) and was therefore confusable with \([\text{ate}^b\text{ma}]\), as \([\text{ate}^b\text{ma}]\) is also expected to be perceived as \([\text{ate}^b\text{ima}]\). In order to probe the illusory vowels in the crucial test words without medial vowels directly, we ran an identification task on the same participants using the same stimuli. The participants were asked to choose the vowel between the two consonants and were given three choices – “i”, “No vowel”, and “Other Vowel”. The instructions and the options were provided in Pinyin for the Mandarin listeners. “Other Vowel” was a choice since there is no unique letter for \([\text{a}]\) in either English orthography or Pinyin.

A mixed ANOVA of the identification rates revealed a main effect of response \([F(2, 74) = 68.58, p < .001]\), an interaction of response by language \([F(2, 74) = 8.64, p < .01]\), an interaction of response by test item \([F(2, 74) = 6.24, p < .001]\), and an interaction of response by test item by language \([F(2, 74) = 4.64, p < .001]\). We ran planned t-tests to compare the identification rates for each response for each of the crucial test items without a medial vowel (Figure 2).

![Identification Rates](image)

**Figure 2:** Exp. 2 – Mean % of Response for each crucial test item (error bars = 1 S.E.; ** = p<0.01)

Results suggest that when a more direct identification task is presented to the Mandarin listeners, only the expected illusory vowels are perceived. This suggests that the unexpected confusions in Exp. 1 could indeed have been due to other perceptual changes.

**Discussion:** Contrary to previous claims [3,4], and in support of our predictions, Mandarin listeners were perceiving different illusory vowels in different phonotactic contexts. Specifically, when presented with phonotactically illegal alveopalatal coda consonants, Mandarin listeners perceived an illusory \([i]\), but in illegal alveolar stop coda contexts, they perceived a \([\text{a}]\). Our results suggest that phonological knowledge is employed not just to identify the locus of a perceptual repair in phonotactically illicit noncewords, but also to infer the illusory vowel itself.

**References**


