

Perception of accent contrasts in vowel devoicing in Japanese

It has been observed in the Japanese literature that when a lexically accented vowel undergoes devoicing, the accent is still perceived by the listeners (e.g., Kitahara & Amano 2001, Maekawa 1990, Sugito & Hirose 1988). Building on this finding, we investigate whether different *types* of contrasts are perceived equally well. We also examine the role of morphological boundaries in vowel devoicing in Japanese. Previous studies have found that the production of vowel devoicing is affected by the morphological boundaries (McCawley 1968, Vance 1992); a question therefore arises as to whether boundaries affect perception as well. In order to investigate these questions, we conducted a perception experiment and found that (i) certain types of accent contrasts are more confusable than others; and (ii) boundaries do not have a consistent effect across all listeners in the perception of accent contrasts involving devoiced vowels.

We test listeners' perception of three accent contrasts taken from Hirayama (2009), who finds that these contrasts are maintained in the production. The contrasts arise from how an accent might resolve if it cannot stay on a devoiced vowel (for details, see Hirayama 2009). The first contrast (hereafter contrast type A) is between items with accent on a devoiced vowel (e.g., *fū'kin* 'nearby'; apostrophes and circles indicate accent and devoicing respectively) and items with accent one syllable after a devoiced vowel (e.g., *fuki'n* 'cloth'). The second contrast (type B) is between items with accent on a devoiced vowel (e.g., *saky'-toka* 'fence-particle') and those with accent on the preceding syllable (e.g., *sa'kyu-toka* 'plan-particle'). The last contrast (type C) is between items with accent on a devoiced vowel (e.g., *kj'kan* 'period') and accentless items (e.g., *kikan* 'trachea').

Our stimuli consisted of 60 minimal pairs differing in terms of accent: 24 pairs each for contrasts A and C and 12 pairs for contrast B. Some of these were control items to set the perception baseline in the non-devoicing environments. The imbalance in number among the contrast types is due to the fact that only contrasts A and C involve boundary effects; we considered boundaries between noun stems and following particles. Eleven speakers (Tokyo accent system) participated in the experiment. When a stimulus was played, they identified the word by choosing one of the relevant minimal pair items given in standard Japanese orthography. The stimuli were randomized for each participant.

The results show that although the contrasts are mostly correctly heard when the items involve devoiced vowels (90.4%, N=1547), not all contrast types are equally well identified: specifically, contrast type A is significantly more confusable (79.7% correct responses, N=562) than the other two contrast types (96.8% for contrast B (N=317); 96.3% for contrast C (N=668)) ($\chi^2(2)=115.174, p < .001$; $p < .001$ for residual analysis). With respect to the effects of the morphological boundaries, there is little or no difference in the rates of the identification whether there is a boundary or not (A: $\chi^2(1)=0.031, p = .861$ (n.s.); C: $\chi^2(1)=3.914, p < .05$). Furthermore, we did not find a consistent boundary effect that holds across all listeners.

We argue that contrast A is more confusable than contrasts B and C because items making contrast A are perceptually less different than those making the other contrasts. We express this by proposing a perception model in which several strong perception cues are matched with the lexical representation of particular accent types. Furthermore, these cues have different weights, in line with Harmonic Grammar (e.g., Smolensky & Legendre 2006, Ramadoss *et al.* 2012); the interaction of the weighted cues in the perception produces different total scores for the items in the contrasts. The discrepancy of scores between contrasting items is smaller in contrast A than in contrasts B and C, and this difference in terms of score discrepancy correlates with a higher degree of confusability for items of this contrast type.