AN ACOUSTIC AND AUDITORY COMPARISON OF IMPLICIT AND EXPLICIT PHONETIC IMITATION

Meaghan Delaney, Soraya Savji, and Molly Babel
Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T 1Z4

1. INTRODUCTION

Talkers have been found to spontaneously accommodate to a model talker in single word shadowing tasks (Babel, 2009, 2010; Goldinger, 1998; Namy et al., 2002; Nielsen, 2008; Shockley et al., 2004). This behaviour has been termed *implicit* phonetic imitation, as talkers have no awareness of having modified their speech. In this study, we compare implicit phonetic imitation to explicit phonetic imitation using an acoustic measure of vocalic phonetic distance and a holistic AXB perceptual similarity measure.

2. METHOD

2.1 Speech Production

Two groups of participants (n = 20, male = 4) completed an auditory naming task which included 50 monosyllabic words with the vowels /i æ ɑ o u/ produced by a male model talker. With the exception of task instructions, the procedure was identical for both groups. In the task instructions the EXPLICIT group was told to explicitly imitate the model talker. The other participants, termed the IMPLICIT group, were instructed to repeat the words naturally. For both groups, the task consisted of a Pre-task reading of the wordlist, followed by a series of three shadowing blocks and a final Post-task reading of the wordlist. This created a total of 800 productions per participant.

2.2 Acoustic Analysis

First and second formants were extracted from the single word productions from a series of Gaussian windows spanning the middle 50% of the vowel with a 2.5 ms step size. Vowel formants for the participants and model talker were normalized using the Lobanov (1971) method.

The Euclidean distance was calculated from each word production to that of the model talker. To calculate the amount of phonetic change in vowel production as a result of exposure to the model talker, the original baseline distance of each word was subtracted from the distance for each following instance of that word for each talker. This created a difference in distance (DID) value. A negative DID value indicates the phonetic distance between the participant and the model talker shrank. A positive value indicates an increase in phonetic distance and phonetic divergence. A value of 0 indicates no change as a result of auditory exposure to the model talker.

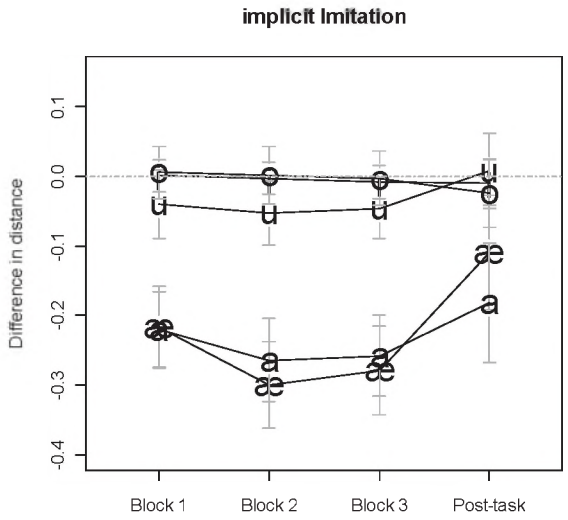

Figure 1. Implicit phonetic imitation by Vowel and Block. The Difference in Distance measure indicates the amount of phonetic imitation. A value of zero shows no change, a negative value indicates imitation and a positive value indicates divergence.

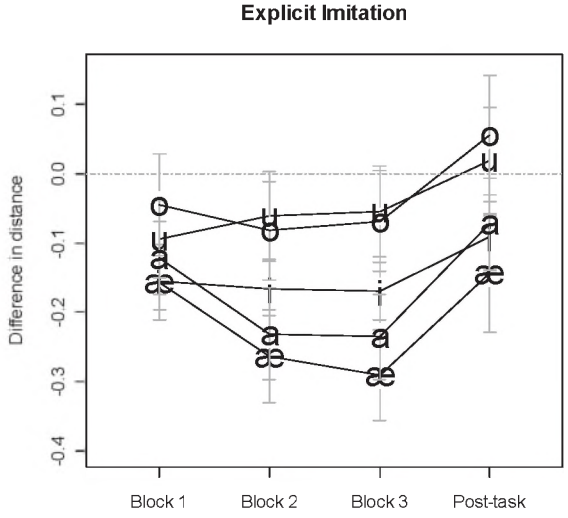

Figure 2. Explicit phonetic imitation by Vowel and Block. The Difference in Distance measure indicates the amount of phonetic imitation. A value of zero shows no change, a negative value indicates imitation and a positive value indicates divergence.

2.3 AXB Perceptual Similarity Task

Participants (n = 13, male = 5) completed an AXB task which included 27 of the monosyllabic words from the

original auditory naming task ending in stop consonants. In the task instructions, participants were told to determine whether the first or third word in a sequence of three identical words sounded more like the middle word. The stimuli consisted of productions from the pre-task, original model talker and the third shadowed block from the auditory naming task. The interstimuli interval was 200ms. The stimuli were blocked by talker and randomized. Counterbalancing was used, whereby participants heard both possible series; pre-task, model talker, shadowed, and shadowed, model talker, pre-task. These responses were analyzed for the proportion of shadowed productions judged more similar to that of the model talker.

3. RESULTS

3.1 Acoustic Analysis

A mixed effects linear model with Subject as a random effect and Block, Vowel, Condition , and Talker Gender as fixed effects was fit to the data using DID as the dependent measure. Figures 1 and 2 highlight some of the main results from the acoustic analysis. In short, the effects of the vowels and blocks across the two conditions were similar. The key exception to this was a main effect of Condition for the /i/ vowel ($\beta = 0.12$, $t = 2.1$), where /i/ was found to be imitated more in the Explicit Condition than in the Implicit.

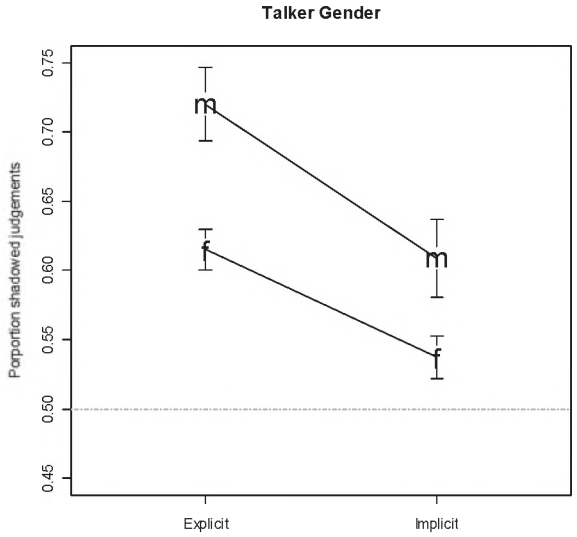

Figure 3. This figure illustrates the main effects of Condition and Talker Gender (male or female). Listeners judged shadowed tokens as more similar to the model talker for talkers in the Explicit condition. Listeners also judged male talkers' shadowed productions as more similar to the model.

3.2 Perceptual Analysis

The proportion of shadowed productions judged as more similar-sounding to the model talker were entered into a mixed effects model with Condition, Talker Gender, and Vowel as fixed effects; Subject was entered as a random effect. In the interest of space, we will report a small selection of the results. There was a main effect of Condition [β =-0.106, t = -5.74, $p < 0.001$] and Talker Gender [$\beta = 0.11$, $t = 3.91$, $p < 0.001$]. These results are shown in Figure 3. There were several significant interactions between Condition and Vowel environment; these are presented in Figure 4.

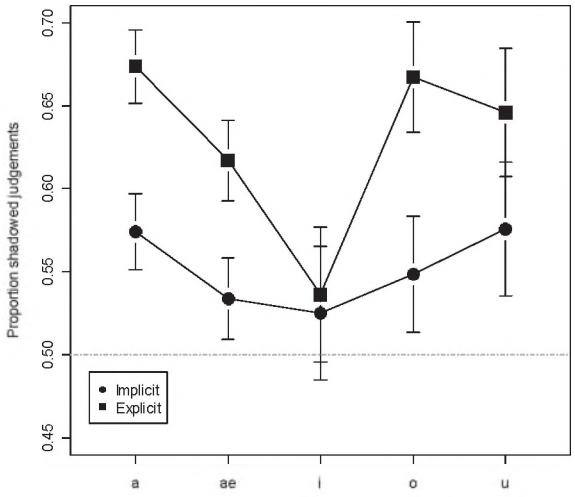

Figure 4. The Condition by Vowel interactions.

4. DISCUSSION

The similarity between explicit and implicit vowel imitation suggest similar cognitive biases mediate the two processes. The similarity judgement task, however, revealed that listeners are more sensitive to explicit imitation. Considering what we found for vowel imitation, this suggests that talkers are imitating more than just formant frequencies and that listeners are sensitive to these differences.

REFERENCES

Babel, M. (2009). *Phonetic and Social Selectivity in Speech Accommodation*. Ph.D. Dissertation, UC Berkeley.

Babel, M. (In press). Dialect divergence and convergence in New Zealand English. *Language in Society,* 39, 1-20.

Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psych. Review*, 105:251–279.

Lobanov, B. (1971) Classification of Russian vowels spoken by different listeners. *JASA*, 49: 606-08

Namy, L., L. Nygaard, & D. Sauterberg. (2002). Gender differences in vocal accommodation: the role of perception. *J of Lang &SocPsych.*, 21:422–432.

Nielsen, K. (2008). *The specificity of allophonic variability and its implications for accounts of speech perception*. Doctoral Dissertation, UCLA.

Shockley, K., L. Sabadini, & C.A. Fowler. (2004). Imitation in shadowing words. *P&P*, 66:422–429.

ACKNOWLEDGEMENTS

Order of the first two authors is alphabetical. Thanks to Keith Johnson for collecting the explicit imitation data at UC Berkeley, Grant McGuire and his lab for help in the data coding, and to UC Berkeley and UBC for financial support. A special thank-you goes out to our research participants as well.