# L2 ENGLISH VOWEL LEARNING BY MANDARIN SPEAKERS: DOES PERCEPTION PRECEDE PRODUCTION?

**Ron I. Thomson**
Dept. of Applied Linguistics, Brock University, St. Catharines, ON, Canada, L2S 3A1, rthomson@brocku.ca

## 1. INTRODUCTION

Many researchers argue that inaccuracy in L2 speech perception is the source of L2 accent (Flege, 1995), while others contend that L2 speech perception and production develop simultaneously (Best, 1995). There is little evidence that L2 production can precede L2 perception. Sheldon & Strange (1982) found that Japanese speakers were able to accurately produce the English /l/-/r/ contrast without being able to perceive it. However, that ability might have resulted from prior pronunciation instruction, coupled with a reading task where speakers could more easily attend to articulation. Munro and Derwing's (2008) study of Mandarin and Slavic speakers' acquisition of English vowels found learners' production ability improved in the absence of explicit pronunciation instruction. Because their research elicited L2 English production using auditory prompts, we can assume that the learners' ability to perceive English vowels must also have improved to at least the same extent.

The purpose of this study is to explicitly investigate the relationship between L2 English vowel identification and production by adult L1 Mandarin immigrants to Canada, using the same auditory stimuli to test both identification and production.

## 2. METHOD

The data analyzed in this experiment were part of a larger study that examined the effect of computer-mediated instruction on the development of L2 English vowels.

### 2.1 Participants

The participants were fourteen female and eight male Standard Mandarin speakers (*M* age = 36.4; range = 27-50) who were new to Canada (*M* LOR = 11.3 months; range = 4-48 months). They had been attending a local ESL program for an average of 4.8 months (range = 1-14 months). Those who had resided in Canada longest reported having had little interaction in English. All reported normal hearing.

2.2 Identification and production tasks

Participants identified English /i/, /ɪ/, /e/, /ɛ/, /æ/, /ɒ/, /ʌ/, /o/, /ʊ/ and /u/ embedded in /bV/ and /pV/ syllables that were produced by two different native speakers, one female and one male. Open syllables were used in order to adhere to Mandarin phonotactic constraints. Participants heard the stimuli via headphones and were asked to click on one of ten nautical flags presented on a computer screen. They had learned to associate these distinctive flags with the English vowel categories during a larger training study. Stimuli used in this study, however, were not used in training.

Productions were elicited using the same /bV/ and /pV/ stimuli described above, this time embedded in the carrier phrase, "The next word is ____." Participants responded by saying, "Now I say ____." Productions were recorded in a quiet room using a Marantz digital recorder with a sampling rate of 44,100 Hz.

### 2.3 Analysis

L2 production responses were acoustically analyzed using F1, F2 and F3 measures extracted from 20% and 70% of each vowel token's duration. Mean F0 and vowel duration were also extracted. These measures were tested against a native speaker English discriminant analysis model (Thomson, 2007) that determined to what English vowel categories the L2 productions were most phonetically similar. Results of this approach are highly correlated with human listener responses (Nearey & Assmann, 1986).

## 3. RESULTS

Mean % correct scores for each vowel on the identification and production tests are shown in Figures 1 and 2, for responses to Voice 1 and Voice 2 stimuli respectively.

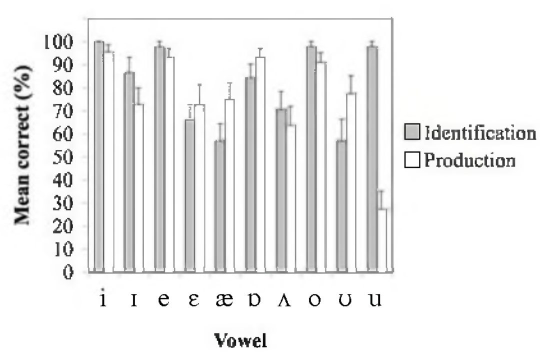


Figure 1. Comparison of average vowel identification and production recognition scores in response to Voice 1 stimuli. Error bars represent standard errors.

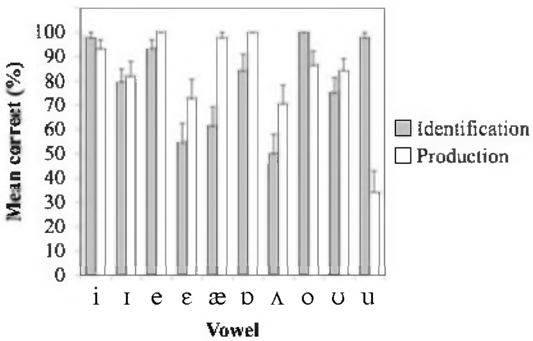


Figure 2. Comparison of average vowel identification and production recognition scores in response to Voice 2 stimuli. Error bars represent standard errors.

A three-way repeated measures analysis of variance was conducted, with Test (2 levels), Stimulus Voice (2 levels) and Vowel (10 levels) as within-speaker factors. Corrected Huynh-Feldt measures are reported because equality of variance could not be assumed for the Vowel variable. The mean scores on the identification and production tests were not significantly different, $F(1, 344) = .300$, $p = .590$, nor were mean scores in response to each stimulus voice, $F(1, 821) = 2.423$, $p = .135$. However, a significant effect for Vowel was found, $F(6.122, 128.561) = 14.966$, $p < .001$. Additionally, the Test x Vowel, $F(6.526, 137.039) = 16.187$, $p < .001$; Test x Voice, $F(1, 21) = 4.886$, $p = .038$; and Voice x Vowel, $F(6.314, 132.602) = 2.30$, $p = .035$ interactions were all significant.

To further investigate the significant Test x Vowel interaction, a series of post-hoc Bonferroni-adjusted *t*-tests were conducted. The mean identification score for /æ/ (59%) was significantly lower than the mean production score (86%), $t(21) = -4.157$, $p < .001$, while the mean identification score for /u/ (98%) was significantly higher than its mean production score (31%), $t(21) = 9.027$, $p < .001$. An examination of error patterns indicated that on the identification test, /æ/ was most often confused with /ɛ/ (20%) followed by /ɒ/ (8%). Conversely, on the production test, /æ/ was most often confused with /ɒ/ (9%), but rarely with /ɛ/ (3%). On the identification test, /u/ was seldom misperceived. In contrast, on the production test /u/ was frequently confused with both /ʊ/ (41%) and /o/ (27%). No other Time x Vowel contrasts were found to be significant.

For the Test x Voice interaction, Bonferroni-adjusted *t*-tests indicated that learner productions in response to Voice 2 stimuli were significantly more accurate (82%) than were productions in response to Voice 1 stimuli (76%), $t(21) = -2{,}968$, $p = .007$. Differences in response to each stimulus voice on the identification test were not significant.

Post-hoc tests examining the Voice x Vowel interaction found no significant differences, likely due to the conservative nature of multiple comparison tests.

## 4. DISCUSSION

The results of this study indicate that the perception and production of most L2 English vowels develop simultaneously. However, the interaction of test type with vowel suggests that the cues learners use for some contrasts lead to different results on each test type. For example, while /u/ and /ʊ/ can be easily identified by their relative duration, lengthening /ʊ/ in production does not result in a recognizably /u/-like pronunciation. Thomson (2007) found that L2 /ʊ/ productions in response to /u/ were twice as long as /ʊ/ productions in response to /ʊ/. The case of /æ/ is less clear. It is surprising that L2 productions of /æ/ were more accurate than were identification responses, despite the fact that the production task also required perception. Some learners might have had difficulty distinguishing between /ɛ/ and /æ/ on the identification test, resulting in a merged category, the distribution of which was more /æ/-like. If this is true, their production responses would be more similar to /æ/, resulting in higher accuracy in production. Further analysis of the data is needed to test this hypothesis.

Finally, the difference found in learner responses to different stimulus voices has important implications. It suggests that the results of such experiments need to be interpreted with caution when only a single stimulus voice is used. It may also suggest that some voices are easier to perceive, and therefore learn from, than are others.

## REFERENCES

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171-204). Timonium, MD: York Press.

Flege, J.E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.) *Speech perception in linguistic experience: Theoretical and methodological issues*. (pp. 229-273). Timonium, MD: York Press.

Munro, M. J., & Derwing, T. M. (2007). A developmental study of Mandarin and Russian speakers' English vowels. *Language Learning, 58*, 479-502.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics, 3,* 243-261.

Thomson, R. I. (2007). *Modeling L1/L2 interactions in the perception and production of English vowels by Mandarin L1 speakers: A training study.* Unpublished Doctoral Dissertation, University of Alberta, Edmonton, AB.

## ACKNOWLEDGEMENTS

The author thanks Terrance Nearey, Murray Munro and Tracey Derwing for their input. The Social Science and Humanities Research Council supported this research.