

ACOUSTIC REALIZATION AND PERCEPTION OF ENGLISH LEXICAL STRESS BY MANDARIN LEARNERS

Yuwen Lai, Joan Sereno, Allard Jongman

Department of Linguistics, University of Kansas, Lawrence KS 66044
 yuwen.lai@gmail.com, sereno@ku.edu, jongman@ku.edu

1. INTRODUCTION

The acquisition of English lexical stress by beginning and advanced Mandarin L2 learners was examined through an acoustic study and perception study. Mandarin Chinese is a tone language which adopts F0 height and contour to signal lexical meaning [1]. English is a stress language which utilizes duration, F0, intensity, and spectral composition to signal stress [2]. Given the suprasegmental differences between the two target languages, Mandarin L2 learners are predicted to differ from native English speakers in the acoustic realization as well as the perception of English lexical stress.

First language transfer occurs when speakers or writers carry over the knowledge from their first language (L1) to a second language (L2) [3]. Most research on L2 acquisition focuses on acquisition at the segmental level, i.e. production and perception of individual consonants or vowels [4]. The current study is one of the few studies which focuses on the phonetic realization and perceptual cue weighting of L2 prosodic acquisition.

2. ACOUSTIC STUDY

2.1 Methods

Fourteen disyllabic word pairs with similar segmental composition but contrasting in stress location (e.g. *contract* and *contract*) were recorded by 18 Mandarin learners of English (9 advanced, 9 beginning). Mean F0, max F0, intensity, duration, and F2 of stressed and unstressed vowels were measured and compared to the production of native English speakers.

2.2 Results

An asymmetry of acoustic correlate realization in nouns and verbs was found in native speakers' production: native speakers utilize mean F0, max F0, intensity, and duration to signal stressed syllables in noun readings, but use only duration in stressed syllables in verb readings. Advanced and beginning Mandarin L2 learners demonstrate a more consistent use of correlates in nouns and verbs: the four correlates are all adopted in both conditions, and the magnitude (difference between stressed and unstressed vowels) was more consistent across nouns and verbs (see Figure 1).

3. PERCEPTION STUDY

3.1 Methods

Fifty Mandarin L2 learners of English (25 beginning, 25 advanced) and 25 native English speakers participated in a stress localization experiment. The stimuli were resynthesized versions of a naturally produced non-word disyllable [dada]. The target perceptual cues tested were max F0 and duration. Spectral composition of the first

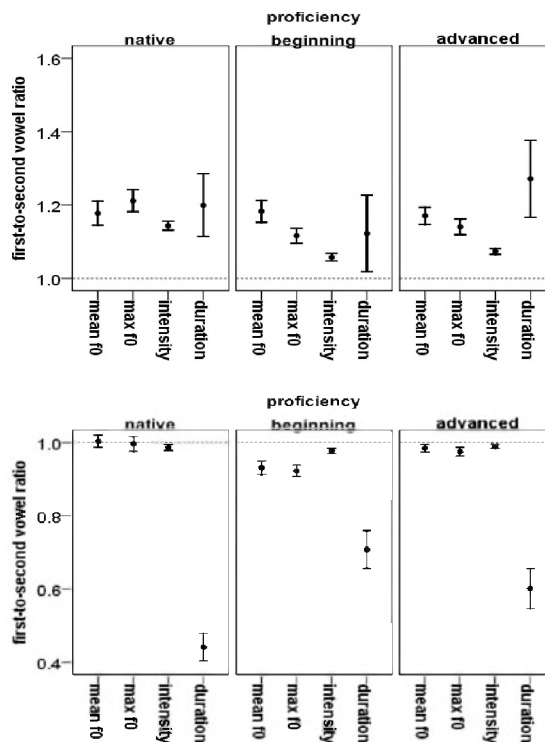


Figure 1. 95% confidence interval of first-to-second vowel ratios for mean F0, max F0, intensity, and duration for nouns (top) and verbs (bottom). The reference line (1.0) is the assumed baseline when a given correlate is identical in both syllables.

vowel of the disyllable was resynthesized to be either full or reduced. Based on the results from the acoustic study, five first-to-second vowel ratios for max F0 and duration were resynthesized for the full and reduced vowel disyllables. The ratios for max F0 are 1.23(F5), 1.115(F4), 1.0(F3), .897(F2), and .813(F1). The ratios for duration are 2.22(D5), 1.11(D4), 1.0(D3), .901(D2), and .45(D1). Participants were instructed to listen to a token and then indicate which syllable was stressed (syllable 1 or syllable 2) by clicking the corresponding text, DAda (first syllable stressed, corresponding to the response '1') or daDA (second syllable stressed, corresponding to the response '2').

3.2 Results

The results indicate that listeners at all proficiencies are sensitive to the spectral composition of the first vowel, i.e., full vowels in the first syllable attracted more noun responses, while reduced vowels in the same location received significantly more verb responses. The interaction of proficiency and max F0 is shown in Figure 2 while the interaction of proficiency and duration is shown in Figure 3. According to the data, native speakers are sensitive to both duration and max F0. Beginning learners are highly

sensitive to duration, but not max F0. Advanced learners' responses are more affected by max F0, and to a much lesser extent by duration.

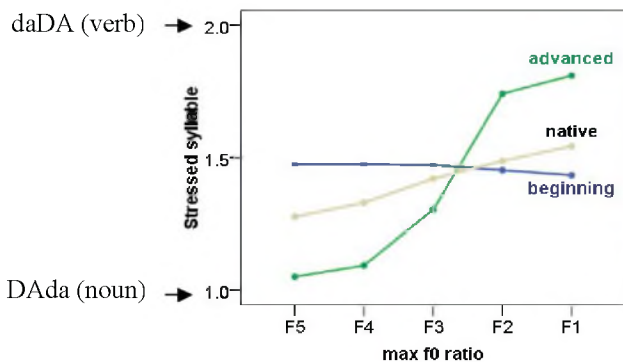


Figure 2. Responses for five max F0 ratios from three participant groups.

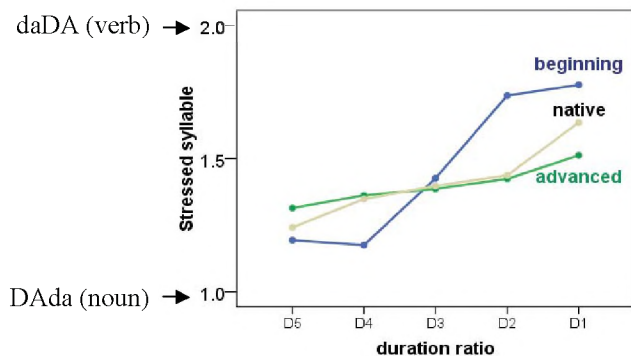


Figure 3. Responses for five duration ratios from three participant groups.

Production and perception of native speakers

Native listeners' responses for each max F0 and duration ratio combination are plotted in Figure 4. The bars represent responses- noun when below 1.5 chance level; verb when above 1.5.

When further compared with the production data, a striking pattern emerges: the role of max F0 is parallel in the production and perception of native speakers. That is, in production, native speakers utilize max F0 to encode stress in nouns but not in verbs. In their perception, the max F0 cue is used in the noun context but not in the verb context. That is, in max F0 cue-only conditions (row D3), native speakers have more noun responses for noun contexts (F5 and F4), but performed at chance level for neutral and verb contexts (F3, F2, and F1).

There is also a correspondence between the production and perception of duration by the native speakers. Native speakers utilized duration in both noun and verb contexts but the magnitude used in verbs was much greater than in nouns. When perceiving a stressed syllable, a greater duration difference is required to trigger verb responses compared to noun responses. In the duration cue-only conditions (column F3), native speakers have more noun responses for noun contexts (D5 and D4), and close to chance levels for neutral and verb contexts (D3, D2, and D1).

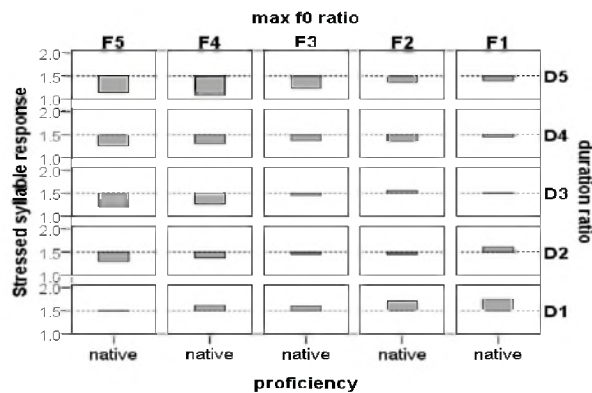


Figure 4. Responses from native English speakers for all max F0 and duration cue combinations.

4. GENERAL DISCUSSION

The current study discovered an asymmetry in correlate realization across noun and verb contexts (*contract* and *contract*) for native speakers. Results from a production and perception study revealed a congruence pattern between the utilization of acoustic correlates and the sensitivity of the perceptual system to a given cue in a specific context. More specifically, the magnitude of a correlate required to obtain stressed syllable perception is parallel to the magnitude of its realization in production.

L2 learners did not show such an asymmetry in their productions in correlate realization across noun and verb contexts. The phonemic feature of F0 in the learners' L1 may constrain them to use F0 in a highly consistent manner. Consequently, beginning learners use a consistent magnitude of F0 across noun and verb contexts. In addition, beginning learners carry over their insensitivity to F0 height [5] due to the phonemic role of F0 in Mandarin. As a result, beginning learners are not sensitive to max F0 changes but rely on duration in identifying stress.

The acoustic realization patterns of advanced learners are between beginning learners and native speakers. It is argued that as the learners' proficiency improves, they are able to approach a more native-like acoustic realization pattern in production. In terms of perception, as their awareness of English stress increases, they are able to utilize F0 (a cue familiar to them in their L1) as the major cue when perceiving English stress.

In sum, the present study demonstrates a context-specific congruence pattern in native production and perception. In addition, L2 learners with a phonemic F0 feature in their L1 are found to encode F0 in stress realization in a more consistent manner across contexts. It seems that L2 learners' perception is affected by the phonemic role of F0 in their L1 but that this effect may change as their proficiency improves.

REFERENCES

- [1] J. M. Howie. (1976). *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge: Cambridge University Press.
- [2] I. Lehiste. (1970). *Suprasegmentals*. Cambridge, MA: MIT.
- [3] T. Odlin. (1989). *Language Transfer: Crosslinguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
- [4] J. E. Flege. (1995). Second-language speech learning: Theory, findings, and problems. In *Speech perception and linguistic experience: Theoretical and methodological issues*, ed. Winifred Strange, 229-273. Baltimore: York Press.
- [5] J. Gandour. (1983). Tone perception in far eastern languages. *J. Phon.* 11:149-175.