

## REALIZATION OF ENGLISH NARROW FOCUS BY L1 ENGLISH AND L1 TAIWAN MANDARIN SPEAKERS

Tanya Visceglia<sup>a</sup>, Chiu-yu Tseng<sup>b</sup>, Chao-yu Su<sup>b</sup> & Chi-feng Huang<sup>b</sup>

<sup>a</sup>Department of Applied English, Ming Chuan University, Taiwan;

<sup>b</sup>Phonetics Lab, Institute of Linguistics, Academia Sinica, Taipei, Taiwan

orlandotaipei@hotmail.com; cytling@sinica.edu.tw

### ABSTRACT

This study compares the realization of English narrow focus by L1 speakers of English and Taiwan Mandarin. Results show that L1 Taiwan Mandarin speakers produce a much smaller increase in average F0 and amplitude for on-focus words and a much smaller decrease in average F0 and amplitude on post-focus words than L1 English speakers do. Moreover, post-focus compression of F0 range and duration, very strongly realized by L1 English speakers, were entirely absent in L2 speakers' production. Failure to perform post-focus compression of F0 range and duration may be attributable to transfer of L1 prosodic patterns. However, transfer cannot account for L2 speakers' weak realization of on-focus F0 range and amplitude expansion. We argue that the weakness of L2 speakers' on-focus/post-focus contrast realization reflects limitations on L2 speech processing, and that weak realization of focus contrasts may also contribute to listeners' difficulty in interpreting the intended focus of L2 utterances.

**Keywords:** L2 prosody, narrow focus, post-focus compression, Taiwan Mandarin, L2 English

### 1. INTRODUCTION

Prosodic realization of utterance-level prominence serves important communicative functions in speech. It can mark the intended focus of a sentence, highlight aspects of information structure by distinguishing between given versus new information, or signal contrastive interpretation, such as the clear difference in speaker's intention between "I don't *think* you stole the money," (but I'm not really sure) and "I don't think *you* stole the money" (but someone else may have).

Often, it is a challenge for L2 speakers to produce the prosodic features that serve to convey such information. Failure to realize prominence and misplacement of prominence have been demonstrated to have a negative effect on L2

speakers' level of comprehensibility [8]. That is to say, L2 speakers' non-target realization of prominence may contribute to listeners' difficulty in extracting their intended meaning or in following their discourse structure [4]. Target-like realization of L2 focus is an important feature for speakers to master; thus, L2 prosody research should further develop our understanding of how the phonetic realization of L2 focus differs from L1, and which of those differences have the greatest impact on comprehensibility.

A growing body of research investigating cross-linguistic differences in the acoustic cues used to realize focus provides a foundation for investigations of L2 focus, allowing us to determine whether L2 differences can be attributed to transfer of first language prosodic strategies, or whether some L2-universal prosodic constraints, such as the tendency of L2 speakers to produce shorter phrase groupings, may also contribute to perceived differences in the realization of prominence [6].

Recent research suggests that Taiwan Mandarin and English represent two different strategies for realization of narrow or contrastive focus: English exhibits on-focus f0 range/amplitude expansion and post-focus f0 range/amplitude compression, and Taiwan Mandarin exhibits on-focus increase in amplitude and duration, but no post-focus compression of f0 range, intensity or duration [2, 9]. Thus, pairing these two languages in a production study would help to investigate the question of whether differences between L1 and L2 realization of narrow focus can be attributed to transfer of L1 prosodic strategies, L2-universal processing strategies, or a combination of the two.

### 2. METHOD

#### 2.1. Materials

Fifteen sets of question/answer pairs were constructed for this experiment. Question sentences were designed to provide a

disambiguating context in order to clarify for participants which word in the answer should receive narrow/contrastive focus. Answers are all in the form of declarative sentences. Each sentence contains one word in contrastive focus, written in boldface all-caps, and one or more words in the post-focus position. Each answer sentence contains a different lexical item in narrow focus; the fifteen items were chosen to represent a range of syllabicities and stress types. Participants were required to produce the answer sentences only, and to stress the word appearing in capital letters. An example appears below:

Disambiguating sentence:

Will 3-day delivery be fast enough?

Experimental sentence:

No. We need **OVERNIGHT** delivery.

## 2.2. Participants

Participants were recruited from university campuses and research institutions in Taiwan. The 8 L1 speakers (4 male, 4 female) are native speakers of North American English. The 9 L2 speakers (5 male, 4 female) are native speakers of Taiwan Mandarin who have received at least ten years of English instruction. Most of the L2 speakers also have some knowledge of Taiwanese, though they are all strongly Mandarin dominant.

## 2.3. Procedure

Speech data were recorded by trained proctors in quiet rooms directly into a laptop computer. Proctors used a recording platform developed specifically for this project, which has pre-loaded experimental sentences, each appearing individually on a computer screen. Participants wore head-mounted Sennheiser PC155 microphones positioned 2 cm away from their mouths, and they were instructed to speak naturally at a normal rate and volume.

## 2.4. Data analysis

255 English utterances (120 L1, 135 L2) were selected for analysis. Speech tokens were sampled at a rate of 16kHz with a quantization of 16 bits. All data were pre-processed for segmental labeling using phone sets from the CMU electronic dictionary, which were then manually spot-checked by trained transcribers for segmental alignment. A trained transcriber additionally labeled two areas in each sentence: on-focus (consisting of the focused word itself) and post-

focus (the words following the focused item up to the next intermediate or intonation phrase break) [1].

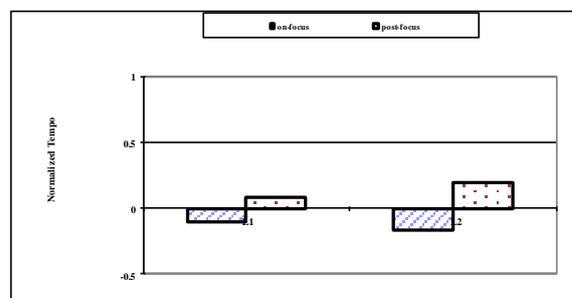
Duration, mean F0, F0 range, and average amplitude were derived for the on-focus and post-focus areas for the purpose of comparison between L1 and L2 speaker groups. Normalization methods were developed to remove features which we believed to be likely to interact with the features under observation, namely inherent differences between L1 and L2 speech rate and pitch range, and the number of phones in the on-focus and post-focus areas. Note that the results presented in Section 3 are represented proportionally, abstracted away from their original units of measurement, in order to highlight differences in the relative salience of the acoustic contrasts produced by the two groups.

## 3. RESULTS

### 3.1. Duration

Figure 1 below shows differences in on-focus expansion and post-focus compression of duration. L2 speakers appear to be slower in the post-focus area. Implications of this finding will be discussed in Section 4.

**Figure 1:** Temporal comparison between on-focus and post-focus words produced by L1 and L2 speakers.



**Table 1:** Duration ANOVA comparison of on-focus and post-focus for L1 and L2 speakers.

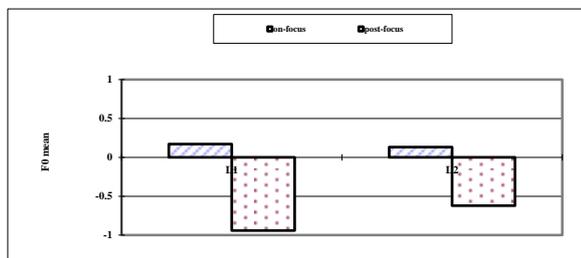
	L1	L2
On-focus	-0.09255	-0.1628
Post-focus	0.09369	0.1984
F-ratio	5.3802	18.58
P-value	0.0213	<0.0001

### 3.2. Mean F0

Figure 2 shows comparison of derived mean F0 between focused and post-focused areas. Although both L1 and L2 speakers contrast average F0 between on-focus and post-focus areas, L1

speakers do so much more robustly. An interpretation for this finding will be offered in Section 4.

**Figure 2:** F0 mean comparison of on-focus and post-focus words produced by L1 and L2 speakers.



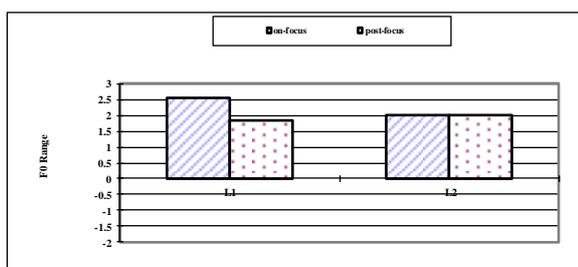
**Table 2:** F0 mean comparison ANOVA of on-focus and post-focus for L1 and L2 speakers.

	L1	L2
On-focus	0.1713	0.1313
Post-focus	-0.9421	-0.6206
F-ratio	209.53	90.174
P-value	<0.0001	<0.0001

### 3.3. F0 range

Measurement of on-focus F0 range expansion and post-focus compression across speaker groups reveals that post-focus compression is entirely absent in the L2 speaker data. Possible interpretations will be discussed in Section 4.

**Figure 3:** F0 range comparison of on-focus and post-focus by L1 and L2 speakers.



**Table 3:** F0 range ANOVA comparison between on-focus and post-focus by L1 and L2 speakers.

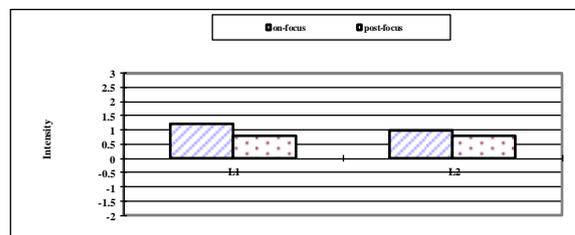
	L1	L2
On-focus	2.528	1.986
Post-focus	1.843	1.995
F-ratio	16.588	0.005934
P-value	<0.0001	0.9387

### 3.4. Amplitude

Comparison of amplitude contrasts between on-focus and post-focus words (see Table 4) yields a pattern similar to the mean F0 data presented in Section 3.2: an amplitude contrast is produced by

both speaker groups. However, it is produced much more robustly by L1 speakers. Section 4 will discuss possible interpretations for this finding.

**Figure 4:** Amplitude comparison of on-focus and post-focus for L1 and L2 speakers.



**Table 4:** Amplitude comparison ANOVA of on-focus and post-focus for L1 and L2 speakers.

	L1	L2
On-focus	1.244	1.013
Post-focus	0.7879	0.8195
F-ratio	144.83	45.988
P-value	<0.0001	<0.0001

## 4. DISCUSSION

### 4.1. Duration

L2 speakers tend to exhibit longer durations in both the focus and post-focus areas. Since these data have been normalized to subtract the effect of between-group differences in speech rate and number of phones, we can reasonably infer that L2 speakers' productions are longer in duration because they are less likely to reduce the unstressed syllables in on-focus words or to reduce words in the post-focus area. L2 speakers' failure to reduce and/or delete de-phrased and unstressed syllables is a well-known characteristic of L2 speech rhythm, particularly for speakers whose L1 is syllable- or mora-timed [3]. Thus, transfer of Taiwan Mandarin's syllable-timed template or focus marking strategy is a likely explanation for this finding, although closer examination of segmental information, such as the vowel quality of the vowels appearing in unstressed syllables, is beyond the scope of this paper.

### 4.2. Mean F0

We see in Figure 2 and Table 2 that although both L1 and L2 speakers contrast average F0 between on-focus and post-focus areas, L1 speakers do so to a much greater extent. This finding cannot be attributed to transfer of L1 prosodic strategy, since Taiwan Mandarin does not use F0 height contrast as a cue to narrow focus. However, weak

realization was also observed in our study of L1 Taiwan Mandarin speakers' production of lexical stress contrasts [7]. Participants were able to acoustically differentiate stressed and unstressed syllables in English multisyllabic words when those words were presented in isolation, but not when they were embedded in higher-level prosodic contexts, i.e. in narrow-focus conditions or at sentence boundaries. We believe that the competing processing demands of simultaneously encoding segmental, lexical stress, and focus information create a processing overload, which weakens realization of the acoustic contrasts used to mark prominence at both the lexical and utterance levels.

### 4.3. F0 range

One of our most striking findings is that post-focus F0 range compression is entirely absent in L2 speakers' data (see Figure 3). We could plausibly attribute this absence to L1 prosodic transfer, as Taiwan Mandarin has been reported to exhibit no post-focus compression. Pitch range is almost identical in the on-focus and post-focus areas for L2 speakers; whereas L1 speakers' pitch range in the on-focus area is 72.9% larger than it is in the post-focus area. Post-focus compression has also been reported to be a highly salient feature in perception of narrow focus by L1 speakers [2].

### 4.4. Amplitude

Comparison of amplitude contrasts between on-focus and post-focus words yields a pattern similar to the mean F0 data discussed in Section 4.2: an amplitude contrast is produced by both speaker groups. However, L1 speakers realize the contrast much more strongly than L2 speakers do. Again, this is not attributable to transfer, since on-focus amplitude expansion has been reported to be a salient component of Taiwan Mandarin's prosodic realization of narrow focus. We believe this finding is also related to the L2 processing difficulty of simultaneously encoding many levels of prosodic information in speech.

## 5. CONCLUSION

In these data, L2 speakers appear to be producing L1-specific prosodic patterns (i.e. absence of post-focus compression) combined with patterns of contrast realization similar to those produced by L1 speakers, but in a weaker form. We believe that weak realization or absence of the prosodic cues

used to realize narrow focus may contribute to the perception that L2 speech is insufficiently differentiated with respect to prominence. We are currently designing a perception study using LPC resynthesis to investigate the relative salience of individual and combined acoustic cues to L1 and L2 listeners' perception of speakers' intended focus. We also plan to compare Taiwan and Beijing Mandarin speakers (Beijing Mandarin, like English, has been reported to exhibit post-focus compression [2]) and to investigate possible interaction effects of sentence position and syllabicity on L2 speakers' production, as these factors have been reported to influence L1 Mandarin listeners' ability to perceive prominence in English [5].

## 6. REFERENCES

- [1] Beckman, M.E., Pierrehumbert, J.B. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3, 255-309.
- [2] Chen, S.-W., Wang, B., Xu, Y. 2009. Closely related languages, different ways of realizing focus. *Interspeech 2009* Brighton, UK.
- [3] Jian, H.L. 2004. On the syllable timing in Taiwan English. *Proceedings of Speech Prosody 2004* Nara.
- [4] Pickering, L. 2004. The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes* 23, 19-43.
- [5] Rosenberg A., Hirschberg, J., Manis, K. 2010. Perception of English prominence by native Mandarin Chinese speakers. *Speech Prosody 2010* Chicago.
- [6] Ueyama, M., Jun, S. 1998 Focus realization in Japanese English and Korean English intonation. *CSLI*. Cambridge Univ. Press, 7, 629-645.
- [7] Visceglia, T., Tseng, C., Su, Z., Huang, C. 2010. Interaction of lexical and sentence prosody in Taiwan L2 English. *SLaTE Workshop, Interspeech 2010* Tokyo.
- [8] Warren, P., Elgort, I., Crabbe, D. 2009. Comprehensibility and prosody ratings for pronunciation software development. *Language Learning & Technology* 13(3), 87-102.
- [9] Xu, Y., Xu, C.X. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33, 159-197.