811.111'342
https://doi.org/10.18485/bells.2022.14.3

Congchao Hua[*]
Shenzhen Technology University, China

PRONOUN PROSODY IN ENGLISH AND MANDARIN AND ITS ACQUISITION BY MANDARIN-SPEAKING ENGLISH LEARNERS[**]

Abstract
This study compared the prosody of pronouns in English and Mandarin in broad focus, narrow focus, and given information, and examined Mandarin-speaking EFL learners' acquisition of pronoun prosody. Ten L1 English speakers, 10 L1 Mandarin speakers, and 60 L1 Mandarin English learners participated in a question-answer reading task. Acoustic analysis revealed that in English pronouns differed from content words in duration in all three information statuses, whereas in Mandarin pronouns differed from content words in maximum intensity in all three information statuses. As regards L2 acquisition of pronoun prosody, auditory and acoustic analyses revealed that the Mandarin-speaking English learners' acquisition of pronoun prosody improved with their English proficiency, yet L1 prosody impacted their phonetic realization of pronouns.

Key words: prosody; pronouns; English; Mandarin; L2 English acquisition

[*] Email address: cchua2-c@my.cityu.edu.hk
[**] This paper was presented at the Sixth Belgrade International Meeting of English Phoneticians (BIMEP 2022), 25–26 March 2022, Faculty of Philology, University of Belgrade.

1. Introduction

In English speech, prominence is important for intelligibility (Jenkins 2000; Hahn 2004; Zielinski 2008) but difficult to learn for second or foreign language (L2) learners (Ramirez Verdugo 2002; Hua & Li 2016). Research has shown that in first-language (L1) English utterances prominence mostly falls on content words, whereas function words are far less likely to take prominence (Altenberg 1987). This tendency, however, is less evident and less consistent in L2 English speech, as they tend to assign prominence indiscriminately to both content words and function words (Juffs 1990; Deterding 2010; Hua & Li 2016, 2019), probably as a result of L1 prosody transfer. Nevertheless, some studies have revealed that L2 English learners from different L1 backgrounds tend to produce similar prominence patterns in their English speech (Grosser 1993; Barlow 1998; Baker 2010). Therefore, to what extent L2 English learners can acquire the prosody of function words and to what extent L1 prosody may influence their acquisition merit detailed investigation. The current study addressed these two issues by comparing the prosody of pronouns in English and Beijing Mandarin in different information statuses and examining L1 Mandarin L2 English learners' acquisition of prosody of pronouns.

2. Literature Review

2.1. Prosody of function words and content words in English and Mandarin

In English, the prosody of function words and that of content words differ considerably in different information statuses, whereas in Mandarin the difference is less distinct.

More specifically, in English speech under broad focus, prominence falls by default on the last content word (Altenberg 1987; Cruttenden 1997; Wells 2006; Roach 2009), which means that function words generally do not take prominence in broad focus. In narrow focus, prosody overruns grammar (Wells 2006), which eliminates the difference between content words and function words: prominence falls on any constituent under focus, be it a content word or a function word. Likewise, in given information, deaccentuation occurs to the constituent expressing information already known, no matter whether it is a content word or a function word.

Mandarin is a non-stress language. In broad focus, there is no clear prominence on any word (Xu 2004), which means that function words and content words are treated in the same way. Like in English, in narrow focus, prominence falls on any constituent under focus, either a content word or a function word; in given information, the constituent that carries shared information is deaccented (Xu 1999), regardless of word class.

To conclude, previous studies seem to suggest that in English the prosody of function words is different from that of content words in broad focus, but not so much in narrow focus or given information; in Mandarin, function words and content words share similar prosodic patterns in all information statuses, that is, broad focus, narrow focus, and given information.

2.2. L1 Mandarin L2 English Learners' Acquisition of Prominence

Studies have revealed some patterns in L1 Mandarin L2 English learners' acquisition of prominence in both phonological and phonetic perspectives. Phonologically, these learners tend to assign equal stress to given information and focus, and they are particularly inclined to assign stress to utterance-final function words such as pronouns and prepositions (Juffs 1990; Wennerstrom 1994, 1998; Deterding 2010). In addition, word class does play a role in these learners' acquisition of prominence, as they are better at assigning prominence to content words than to function words (as under narrow focus), and better at deaccenting function words than content words (as in given information) (Hua 2021).

Phonetically, the findings are not consistent. For focus, some (McGory 1997; Wennerstrom 1998) report that L1 Mandarin L2 English learners produce L1 English-like duration and intensity, but not F0, while others report that these learners tend to produce shorter duration and lower F0 than L1 English speakers (e.g., Hua & Li 2019). In addition, Barlow (1998) reports that L1 Mandarin English learners rely heavily on intensity when realizing focus.

Studies on these learners' realization of given information have also yielded inconsistent results. Baker (2010) found that L1 Mandarin English learners produce reduced duration for given information, yet the duration they produce still differs from that produced by L1 English speakers. However, Hua and Li (2019) found that these learners can produce duration and f0 comparable to those of L1 English speakers.

In addition, it has been found that these learners' acquisition of sentence prosody improves with their English proficiency (Barlow 1998; Baker 2010; Hua & Li 2019; Hua 2022), yet in an unbalanced way. Their acquisition of the phonological features (i.e., prominence/deaccentuation placement) improves faster and more considerably than their acquisition of the phonetic features (i.e., the phonetic realization of prominence/deaccentuation), the latter being more impacted by L1 prosody transfer and resistant to improvement with proficiency, and their acquisition for focus improves better than that for given information (Hua & Li 2019; Hua 2021, 2022).

Since English and Mandarin differ in their prosody of function words (especially in broad focus), and word class affects L2 acquisition of prominence and deaccentuation, it would be of theoretical and practical values to know if L2 English learners treat function words and content words differently in their English speech, and if there are effects of L1 transfer. The two research questions addressed are as follows:

1) How are pronouns phonetically realized (as compared with content words) in broad focus, narrow focus and given information in English and Mandarin?
   Hypothesis 1: in English, pronouns are realized with weaker phonetic cues than content words in broad focus, but comparable phonetic cues to content words in narrow focus and given information; in Mandarin, pronouns and content words are realized with comparable phonetic cues in all three information statuses.
2) What patterns do L1 Mandarin L2 English learners follow in acquiring prosody of pronouns in different information statuses?
   Hypothesis 2: phonologically, these learners treat pronouns like content words, and their performance approaches L1 English speakers as their proficiency increases.
   Hypothesis 3: phonetically, there is an effect of L1 transfer, which decreases as the learners' proficiency increases.

3. Research Method

This is an experimental study comprising two experiments, both in the form of read-aloud tasks.

3.1. Participants

The participants in Experiment 1 were 10 L1 British English speakers (aged 19–28, five male, five female) and six L1 Beijing Mandarin speakers (aged 23–33, all female). The L1 English speakers were international students and teachers in a city in central China. None of them could speak Mandarin Chinese. The L1 Mandarin speakers were all from northern China with a B.A., M.A. or doctorate degree in humanities. Since English is a compulsory course in the Chinese educational system, all of them could speak English and some were rather fluent. However, they did not use English on a daily basis and Beijing Mandarin is the primary language in their environment.

The participants in Experiment 2 were 60 L1 Mandarin undergraduates from a key university in central China. They were aged 18-22, 30 male and 30 female, having little experience living in English-speaking countries/areas. They were divided into three proficiency groups according to a Cambridge English proficiency test (https://www.cambridgeenglish.org/test-your-english/general-english/): preliminary (8-16 points), intermediate (17-19 points), and advanced (20–25 points). The 10 L1 English speakers in Experiment 1 served as a reference group in Experiment 2.

3.2. Stimuli

The English stimuli were 12 question-answer pairs (4 pairs for each information status, 2 ending with a pronoun and two with a content word). All words in the questions and answers are within the 2000 range of the New General Service List (Browne et al. 2013), and all words in the answers are monosyllables. The following are some samples, with the target words in italics:

(1) Broad focus in English

A: What did you say?	A: What did you say?
B: Dan loves to work with *Jim*.	B: Dan loves to work with *him*.

(2) Narrow focus in English

A: Who does Dan love to work with?	A: Who does Dan love to work with?
B: Dan loves to work with *Jim*.	B: Dan loves to work with *him*.

(3) Given information in English

A: Does Dan love to work without Jim?	A: Does Dan love to work without Jim?
B: Dan loves to work with *Jim*.	B: Dan loves to work with *him*.

The Mandarin stimuli were also 12 question-answer pairs (4 pairs for each information status, 2 ending with a pronoun and two with a content word). All words in the answers are with the first tone (the high level tone), and all target words are monosyllables. The following are some samples, the underlined being the target words:

(4) Broad focus in Mandarin

A: 你 说 什 么?	A: 你说什么?
you say what	
'What did you say'	
B: 张 欢 天 天 喝 汤。	B: 张 欢 天 天 喝它。
Zhang Huan every day drink soup	Zhang Huan every day drink it
'Zhang Huan drinks soup every day'	'Zhang Huan drinks it every day'

(5) Narrow focus in Mandarin

A: 张 欢 天 天 喝 什么?	A: 张欢天天喝什么?
Zhang Huan every day drink what	
'What does Zhang Huan drink every day'	
B: 张 欢 天 天 喝 汤。	B: 张 欢 天 天 喝 它。
Zhang Huan every day drink soup	Zhang Huan every day drink it
'Zhang Huan drinks soup every day'	'Zhang Huan drinks it every day'

(6) Given information in Mandarin

A: 张 欢 天 天 做 汤 吗?	A: 张欢天天做汤吗?
Zhang Huan every day cook soup question marker	
'Does Zhang Huan cook soup every day'	
B: 张 欢 天 天 喝 汤。	B: 张 欢 天 天 喝它。
Zhang Huan every day drink soup	Zhang Huan every day drink it
'Zhang Huan drinks soup every day'	'Zhang Huan drinks it every day'

3.3. Procedure

The recordings from the 10 L1 English speakers and the 60 learners were collected in person before COVID-19, and after COVID-19 broke out, recordings from the six L1 Mandarin speakers were collected through the internet.

The L1 English speakers and the learners were recorded individually with a Sony linear PCM recorder (model PCM-D100) in a quiet room. After receiving the instructions and doing a few trial items on a computer, the participants read the question-answer pairs from the computer screen, the recorder one meter away from their mouth. The L1 Mandarin speakers recorded their reading of the question-answer pairs on their own phones. To do this, they received a pdf file containing the instructions and the stimuli. All three groups were asked to read the question-answer pairs twice at their own pace.

Although difference in recording method may affect the quality and comparability of recordings, diverse recording methods can still yield reliable data for analyzing prosody (Guan & Li 2021). Additionally, in this study raw values of phonetic cues were all converted to ratios, which may further enhance the reliability of the data.

3.4. Data Analysis

The recordings from Experiment 1 were analyzed acoustically using Praat (Boersma & Weenink 2021), and the recordings from Experiment 2 were analyzed both auditorily and acoustically.

The recordings of the answers from the L1 speakers in Experiment 1 were firstly selected. Mostly it was the first readings that were picked. In case where a first reading was disfluent or contained a mispronunciation, the second reading was chosen instead. In total 120 (12 sentences x 10 speakers) were selected. These recordings were checked in Praat, and those with improper displays of F0 and/or intensity were discarded. Altogether 111 sentence recordings from the L1 English speakers and 70 from the L1 Mandarin speakers entered the acoustic analysis. These recordings were annotated manually in Praat, and then values of duration, maximum F0, average F0, and maximum intensity were extracted using Praat scripts. These parameters were selected because Breen et al. (2010) report that they are the best indicators of information status in English. The values

of the four parameters were then converted to ratios by dividing the raw value of the target word by the average value of all words in the same sentence.

The recordings of the answers from the 60 learners in Experiment 2 were also firstly selected in the same fashion as mentioned above, and 720 sentence recordings (12 sentences x 60 learners) entered auditory analysis. These 720 learner recordings, together with the 120 L1 English speaker recordings from Experiment 1, were analyzed auditorily by a trained phonetician twice, assigning 1 to a sentence recording with correct placement of prominence (focus sentences) or deaccentuation (given information sentences) and 0 to a sentence recording with unclear or wrong placement of prominence or deaccentuation. The intra-rater reliability reached .98.

For the acoustic analysis, recordings of sentences with correct placement of prominence or deaccentuation were selected and then checked in Praat. Those with improper displays of F0 and/or intensity were discarded. Thus, 410 sentence recordings from the learners entered the acoustic analysis. These recordings were annotated and ratios calculated in the same way as the L1 recordings.

Then all ratios and scores were analyzed statistically using factorial ANOVA tests in SPSS (26.0).

4. Results and Discussion

4.1. Pronoun Prosody in L1 English vs. L1 Mandarin

The phonetic realizations of pronouns (as compared with content words) in L1 English and L1 Mandarin are presented in Figure 1.

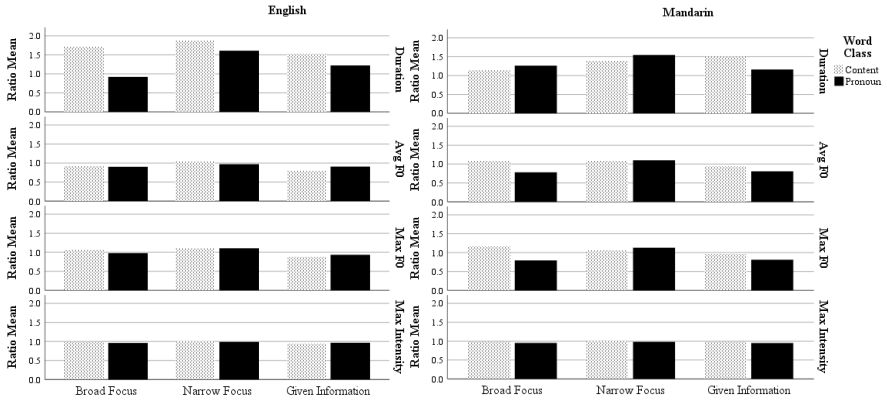

Figure 1. Phonetic realizations of pronouns in English and Mandarin

For English, 2 x 2 x 3 ANOVAs were computed, with ratios as dependent variables and gender (male vs. female), word class (content words vs. pronouns) and information status (broad focus, narrow focus, given information) as independent variables. Results (Table 1a) revealed that duration ratio varied by word class and by information status, and there was an interaction between word class and information status. Post-hoc tests with Bonferroni correction showed that the duration ratio for narrow focus was higher than those for broad focus and given information (both at $p < .001$), and that pronouns were realized with lower duration ratios than content words in broad focus, narrow focus, and given information (all at $p < .001$).

Average F0 ratio and maximum F0 ratio both differed by information status only, with higher values for narrow focus than for given information ($p = .021$; $p = .011$).

Maximum intensity ratio differed by word class and by information status, and there was an interaction between word class and information status. Post-hoc tests showed that the maximum intensity ratio for narrow focus was higher than those for broad focus ($p = .003$) and given information ($p < .001$), and that pronouns were realized with lower maximum intensity ratios than content words in broad focus ($p < .001$) and narrow focus ($p = .009$).

Table 1. Differences indicated by ANOVA results for L1 English and L1 Mandarin

a. L1 English

Duration Ratio	***df***	***F***	***p***	**partial η^2**
word class	1, 105	141.03	< .001***	.57
information status	2, 105	49.77	< .001***	.49
word class*information status	2, 105	18.91	< .001***	.27
Average F0 Ratio	***df***	***F***	***p***	**partial η^2**
information status	2, 105	3.94	.022*	.07
Maximum F0 Ratio	***df***	***F***	***p***	**partial η^2**
information status	2, 105	4.56	.013*	.08
Maximum Intensity Ratio	***df***	***F***	***p***	**partial η^2**
word class	1, 105	8.07	.005**	.07
information status	2, 105	31.75	<.001***	.38
word class*information status	2, 105	11.63	<.001***	.18

b. L1 Mandarin

Duration Ratio	***df***	***F***	***p***	**partial η^2**
information status	2, 64	9.17	< .001***	.22
word class*information status	2, 64	9.61	< .001***	.23
Average F0 Ratio	***df***	***F***	***p***	**partial η^2**
word class	1, 64	15.78	< .001***	.20
information status	2, 64	13.04	< .001***	.29
word class *information status	2, 64	7.36	.001**	.19
Maximum F0 Ratio	***df***	***F***	***p***	**partial η^2**
word class	1, 64	14.09	< .001***	.18
information status	2, 64	9.90	< .001***	.24
word class *information status	2, 64	10.03	< .001***	.24
Maximum Intensity Ratio	***df***	***F***	***p***	**partial η^2**
word class	1, 64	15.72	< .001***	.20
information status	2, 64	3.81	.027*	.11

*significant at .05 level, **significant at .01 level, ***significant at .001 level

For Mandarin, 2 x 3 ANOVAs were run, with ratios as dependent variables and word class (content words vs. pronouns) and information status (broad focus, narrow focus, given information) as independent variables. Results (Table 1b) showed that duration ratio differed by information status, and there was an interaction between information status and word class. Post-hoc tests revealed that narrow focus was realized with a higher duration ratio than broad focus ($p < .001$) and given information ($p = .043$), and that pronouns were realized with a lower duration ratio than content words in given information ($p = .001$).

Average F0 ratio and maximum F0 ratio both differed by word class and by information status, and there were interactions between word class and information status for both. Post-hoc tests showed that narrow focus was realized with higher average and maximum F0 ratios than broad focus ($p = .002$, $p = .037$) and given information (both at $p < .001$), and that pronouns were realized with lower average and maximum F0 ratios (both at $p < .001$) than content words in broad focus.

Maximum intensity ratio differed by word class and information status. Post-hoc tests showed that narrow focus was realized with a higher maximum intensity ratio than given information ($p = .010$).

Table 2. Phonetic realizations of pronouns in L1 English and L1 Mandarin

L1 English				
	Duration	**Average F0**	**Maximum F0**	**Maximum Intensity**
Broad	P<C***			P<C***
Narrow	P<C***			P<C**
Given Info.	P<C***			
L1 Mandarin				
	Duration	Average F0	Maximum F0	Maximum Intensity
Broad		P<C***	P<C***	P<C***
Narrow				P<C***
Given Info.	P<C**			P<C***

P=pronouns, C=content words

**significant at .01 level, ***significant at .001 level

Based on the above results, the phonetic realizations of pronouns in English and in Mandarin are summarized in Table 2. In both languages, pronouns were realized with weaker phonetic cues than content words across the information statuses. This difference, however, is not consistent across the four parameters: the pronouns were not realized with lower values of all four parameters than content words in any of the three information statuses. Thus, Hypothesis 1 that pronouns are realized with weaker phonetic cues than content words only in broad focus in English is rejected.

It seems that although phonologically prosody overrun word class in narrow focus and given information in English, as claimed by Wells (2006), phonetically pronouns were realized with weaker cues than content words in all three information statuses. Similarly, in Mandarin, pronouns were also realized with weaker phonetic cues than content words in all three information statuses. One possible explanation is that in both languages, pronouns, compared with content words, represent known or shared information and are thus less stressed than content words in all information statuses.

Despite the above similarity, there was indeed a major difference between English and Mandarin. In English, pronouns and content words differed by duration in all information statuses, while in Mandarin, they differed by maximum intensity in all information statuses. In both languages, pronouns were realized with lower values in duration (English) or maximum intensity (Mandarin) than content words. Thus, it seems that in English duration is the major phonetic cue distinguishing between pronouns and content words, while in Mandarin it is maximum intensity that distinguishes between the two word classes.

4.2. L1 Mandarin L2 English learners' acquisition of pronoun prosody

As mentioned in Section 3.4, the learners' recordings were analyzed both auditorily and acoustically.

4.2.1. Results of auditory analysis

The auditory analysis showed that these learners performed differently on pronouns and content words (Figure 2).

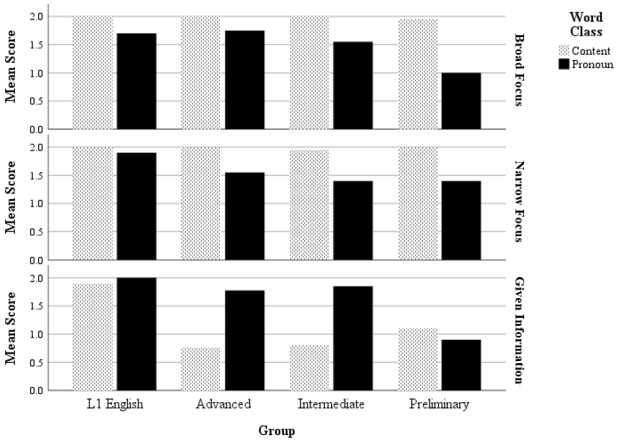

Figure 2. Score by group, information status, and word class

A 2 x 3 x 3 x 2 ANOVA test was run, with score as dependent variable, gender (male vs. female) and group (L1 English, advanced, intermediate, preliminary) as between-subjects independent variables, and information status (broad focus, narrow focus, given information) and word class (content word, pronoun) as within-subjects independent variables. Results (Table 3) revealed differences by group, by information status, and by word class, and interactions between group and word class, between group and information status, between word class and information status, and between group, information status, and word class.

Table 3. Differences indicated by ANOVA results for L2 English: Auditory analysis

	df	***F***	***p***	**partial η^2**
Group	3, 396	16.45	< .001***	.11
Word Class	1, 396	8.11	.005**	.02
Information Status	2, 396	26.14	< .001***	.12
Group*Word Class	3, 396	12.28	< .001***	.09
Group*Information Status	6, 396	4.12	.001**	.06
Word Class*Information Status	2, 396	41.85	< .001***	.17
Group*Word Class*Info. Status	6, 396	4.82	< .001***	.07

**significant at .01 level, ***significant at .001 level

Post-hoc tests showed that the L1 English group outperformed all three learner groups ($p = .002$, $p < .001$, $p < .001$), and that all groups performed better on broad focus and narrow focus than on given information (both at $p < .001$). In addition, while all the other three groups performed comparably on pronouns and content words, the preliminary group performed worse on pronouns than on content words ($p < .001$). While the L1 speaker group performed comparably on all information statuses, the three learner groups all performed worse on given information than on broad focus and narrow focus ($p = .037$- $p < .001$). Moreover, the learner groups generally performed worse on pronouns than on content words in focus (all at $p < .001$), but better on pronouns than on content words in given information (all at $p < .001$). More specifically, the advanced and intermediate groups followed this general pattern ($p = .031$- $p < .001$), but the preliminary group performed comparably on pronouns and content words in given information.

The above results suggest that the L1 Mandarin L2 English learners performed worse than the L1 speakers even at the advanced level, yet the learners did improve as their proficiency increased, implying that they were gradually learning to treat pronouns as L1 English speakers did. Thus, Hypothesis 2 that phonologically the learners treat pronouns like content words, but their performance approaches L1 English speakers as their proficiency increases is retained.

This result is consistent with previous findings that acquisition of L2 prosody improves with L2 proficiency (Baker 2010; Hua & Li 2016, 2019; Hua 2022). However, it is worth noting that these learners were not good at accenting pronouns for narrow focus or deaccenting pronouns for broad focus and given information, and they were especially poor at deaccenting content words for given information. This echoes with previous findings that L1 Mandarin L2 English learners tend to treat given information as focus (Hua 2021) and stress the last word in an utterance regardless of its word class (Juffs 1990; Wennerstrom 1994, 1998; Deterding 2010; Hua & Li 2016, 2019).

These learners' failure in deaccenting pronouns for broad focus and given information suggests a certain degree of L1 transfer, as pronouns and content words tend to be treated equally in their L1 Mandarin. Their failure in accenting pronouns for narrow focus, however, suggests the contrary. Perhaps these learners neutralized pronouns across all information statuses so that the pronouns they produced were neither prominent in narrow focus nor weak in given information compared with L1 English.

4.2.2. Results of acoustic analysis

The results of the acoustic analysis are presented in Figure 3. Four 2 x 3 x 3 x 2 ANOVA tests were run, with ratios as dependent variables, gender (male vs. female) and group (L1 English, advanced, intermediate, preliminary) as between-subjects independent variables, and information status (broad focus, narrow focus, given information) and word class (content word, pronoun) as within-subjects independent variables.

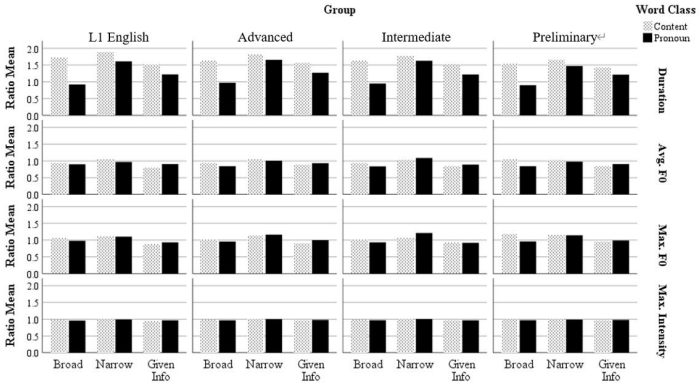

Figure 3. Phonetic realization of pronouns by group, information status, and word class

Results (Table 4) showed that duration ratio differed by group, by information status, and by word class, and there was an interaction between word class and information status. Post-hoc tests revealed that the L1 English group and the advanced group produced higher duration ratios than the preliminary group (p = .006, p =.007), the duration ratio for narrow focus was higher than those for broad focus and given information (both at p < .001), and the duration ratio for given information was higher than that for broad focus (p = .011). Moreover, in all three information statuses, duration ratios for pronouns were lower than for content words (all at p < .001).

Average F0 ratio and maximum F0 ratio differed by information status, and there were interactions between word class and information status for both (See Table 4). Post-hoc tests revealed that the F0 ratios for narrow focus were higher than those for broad focus and given information (all at p < .001), and that in broad focus F0 ratios were lower for pronouns than for content words (p < .001, p = .004).

Maximum intensity ratio differed by group, by information status, and by word class, and there were interactions between group and information status and between word class and information status (See Table 4). Post-hoc tests showed that the L1 English group produced a lower ratio than all three learner groups ($p = .001$, $p = .029$, $p = .002$), that the ratio for narrow focus was higher than those for broad focus and given information, and the ratio for broad focus was higher than that for given information (all at $p < .001$). In addition, in broad focus and narrow focus, pronouns were realized with a lower maximum intensity ratio than content words (both at $p < .001$), but in given information, pronouns were realized with a higher maximum intensity ratio than content words ($p = .011$).

Table 4. Differences indicated by ANOVA results for L2 English: Acoustic analysis

Duration Ratio	***df***	***F***	***p***	**partial η^2**
group	3, 497	7.66	< .001***	.04
word class	1, 497	362.90	.005**	.42
information status	2, 497	154.15	< .001***	.38
word class*information status	2, 497	59.15	< .001***	.19
Average F0 Ratio	***df***	***F***	***P***	**partial η^2**
information status	2, 497	27.28	< .001***	.10
word class*information status	2, 497	7.66	.001**	.03
Maximum F0 Ratio	***df***	***F***	***P***	**partial η^2**
information status	2, 497	27.98	< .001***	.10
word class*information status	2, 497	5.59	.004**	.02
Maximum Intensity Ratio	***df***	***F***	***P***	**partial η^2**
group	3, 497	4.74	.003**	.02
word class	1, 497	22.72	< .001***	.04
information status	2, 497	102.22	< .001***	.29
group*information status	6, 497	2.44	.025*	.03
word class*information status	2, 497	31.43	< .001***	.11

*significant at .05 level, **significant at .01 level, ***significant at .001 level

The learners' phonetic realization of pronouns is presented in Table 5, together with those in L1 English and L1 Mandarin for the convenience of comparison.

Table 5. Phonetic realizations of pronouns in L1 English, L1 Mandarin, and L2 English

L1 English				
	Duration	**Average F0**	**Maximum F0**	**Maximum Intensity**
Broad	P<C***			P<C***
Narrow	P<C***			P<C**
Given Info.	P<C***			
L1 Mandarin				
	Duration	**Average F0**	**Maximum F0**	**Maximum Intensity**
Broad		P<C***	P<C***	P<C***
Narrow				P<C***
Given Info.	P<C**			P<C***
L2 English				
	Duration	**Average F0**	**Maximum F0**	**Maximum Intensity**
Broad	P<C***	P<C***	P<C**	P<C***
Narrow	P<C***			P<C***
Given Info.	P<C***			P>C*

P=pronouns, C=content words

*significant at .05 level, **significant at .01 level, ***significant at .001 level

Table 5 shows that the learners' phonetic realization of pronouns entails features from both English and Mandarin. Like L1 English speakers, they produced shorter duration for pronouns than for content words in all three information statuses. Like L1 Mandarin speakers, they produced lower average F0 and maximum F0 for pronouns than for content words in broad focus. Also, they produced lower maximum intensity for pronouns than for content words in broad focus and narrow focus, which is consistent

with both L1 English and L1 Mandarin. What is tricky is that in given information they produced higher maximum intensity for pronouns than for content words, which is a pattern absent in both L1 English and L1 Mandarin. In general, there is clear impact of L1 prosody on these learners' phonetic realization of pronouns.

In addition, overall the learners produced duration comparable to L1 English speakers at the intermediate stage, average and maximum F0 comparable to L1 English speakers at the preliminary stage, but higher maximum intensity than L1 English speakers even at the advanced stage. Thus, it can be concluded that while duration improved with proficiency, average and maximum F0 and maximum intensity improved very little. Therefore, Hypothesis 3, which states that phonetically there is an effect of L1 transfer, but the effect decreases as proficiency increases, is generally retained.

The findings above are consistent with Barlow (1998), who reports that Mandarin-speaking English learners rely heavily on intensity when realizing focus, but partially consistent with McGory (1997), who report that Mandarin-speaking English learners tend to produce L1 English-like duration and intensity, but not F0. In this study, however, the participants produced duration and F0 comparable to L1 English speakers at a certain stage, but not intensity. This is probably because the proficiency levels of the participants in these two studies are not comparable. Moreover, the finding that the learners' phonetic realization of pronouns improved with their proficiency echoes with Barlow (1998), Baker (2010), Hua and Li (2019) and Hua (2022), all reporting improvement in phonetic realization of L2 English prosody with English proficiency.

However, neither L1 English nor t L1 Mandarin could explain the higher intensity for pronouns than for content words in the learners' phonetic realization of given information, as this pattern is absent in both languages. This unique pattern could be evidence of these learners' struggle with the prosody of given information, which is especially difficult for these learners and even at the advanced level many of them still fail to mark given information with proper prosody in terms of both deaccentuation and phonetic realizaiton (Hua & Li 2019).

Another intriguing point is that duration approached L1 English as these learners' proficiency increased, but average F0 and maximum F0 and maximum intensity improved very little with proficiency. While the improvement in duration shows a sign of learning as expected, the lack of

improvement in maximum intensity could have resulted from the learners' deep-rooted L1 Mandarin prosody, as in L1 Mandarin maximum intensity plays an important role in the realization of information status. The lack of improvement in average F0 and maximum F0, on the other hand, can be attributed to a ceiling effect. As indicated in Table 4, there were no differences in F0 ratios by group, which suggests that these learners could already produce L1 English-like F0 at the early stage of learning. Consequently, there was little room for improvement in F0.

Taken all findings together, while there were signs of learning in these learners, L1 prosody transfer seems difficult to conquer for further improvement.

5. Conclusion

This study compared the prosody of pronouns in English and Mandarin and investigated the acquisition of English pronoun prosody by L1 Mandarin speakers. The major findings are: 1) in both English and Mandarin, the prosody of pronouns does not differ by information status, and pronouns and content words differ in their prosody in all three information statuses; 2) the L1 Mandarin L2 English learners' acquisition of pronoun prosody improves with their proficiency, but there is clear evidence of L1 prosody transfer, especially in phonetic realization of pronouns.

Based on the above findings, it can be concluded that word class is a non-negligible factor in researching and teaching prosody. In practice, it is necessary to raise L2 English learners' awareness of pronoun prosody. For example, it could be made explicit when and how pronouns should be accented or deaccented.

This study is not exempt from limitations. To begin with, male L1 Mandarin speakers should also be included, and it is preferable that all participants should be recorded under the same condition. In addition, it would be more informative if spontaneous speech is also investigated. Further research can address these issues and dive deeper into the prosody-grammar interface in L2 prosody acquisition.

References

Altenberg, B. (1987). *Prosodic Patterns in Spoken English: Studies in the Correlation between Prosody and Grammar for Text-to-speech Conversion*. Lund University Press.

Baker, R. E. (2010*). The Acquisition of English Focus Marking by Non-native Speakers* (Unpublished PhD dissertation). Northwestern University, Evanston.

Barlow, J. S. (1998). *Intonation and Second Language Acquisition: A Study of the Acquisition of English Intonation by Speakers of Other Languages* (Unpublished PhD dissertation). University of Hull, Hull, U.K.

Boersma, P. and D. Weenink. (2020). Praat: Doing Phonetics by Computer [Computer program]. Version 6.1.13. (20 April 2020) <http://www.praat.org>

Breen, M. et al. (2010). Acoustic Correlates of Information Structure. *Language and Cognitive Processes,* 25(7), 1044–1098.

Browne, C. et al. (2013). The New General Service List. (1 March 2019) <http://www.newgeneralservicelist.org>

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge: Cambridge University Press.

Deterding, D. (2010). ELF-based Pronunciation Teaching in China. *Chinese Journal of Applied Linguistics,* 33(6), 3–15.

Grosser, W. (1993). Aspects of Intonational L2 Acquisition. In: B. Kettemann and W. Wieden (eds.), *Current Issues in European Second Language Acquisition Research*, Gunter Narr Verlag Tubingen, 81–94.

Hahn, L. D. (2004). Primary Stress and Intelligibility: Research to Motivate the Teaching of Suprasegmentals. *TESOL Quarterly*, 38(2), 201–223.

Hua, C. C. (2021, December). Prominence for Focus Produced by Mandarin-speaking EFL Learners. Paper presented at Accents 2021, Lordz, Poland/online.

Hua, C. C. (2022, in print). Acquisition of Prosody of Given Information by Mandarin-speaking EFL Learners. In: J. Levis and A. Guskaroska (eds.), *Proceedings of the Pronunciation in Second Language Learning and Teaching (PSLLT) 2021 Conference*.

Hua, C. C. and B. Li. (2016). Nuclear Stress Patterns in Reading by Adult Chinese EFL Learners: Explicit Training or Implicit Learning? *The Journal of Asia TEFL,* 13(1), 1–15.

Hua, C. C. and B. Li. (2019, May). Acquisition of Prosodic Marking of Information Status by Mandarin-speaking Learners of English: Phonological vs. Phonetic Properties. Paper presented at the 6th International Conference on English Pronunciation: Issues and Practices, Skopje, North Macedonia.

Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.

Juffs, A. (1990). Tone, Syllable Structure and Interlanguage Phonology: Chinese Learners' Stress Errors. *International Review of Applied Linguistics*, 28, 99–117.

Guan, Y. H. and B. Li. (2021). Usability and Practicality of Speech Recording by Mobile Phones for Phonetic Analysis. *Proceedings of the 12th International Symposium on Chinese Spoken Language Processing (ISCSLP)* (pp. 1-5). (10 March 2022) <https://ieeexplore.ieee.org/document/9362082>

McGory, J. T. (1997). *Acquisition of Intonational Prominence in English by Seoul Korean and Mandarin Chinese Speakers* (Unpublished PhD dissertation). Iowa State University, Ames.

Ramírez Vergugo, D. (2002). Non-native Interlanguage Intonation Systems: A Study Based on a Computerized Corpus of Spanish Learners of English. *ICAME Journal,* 26, 115–132.

Roach, P. (2009). *English Phonetics and Phonology: A Practical Course* (4th ed.). Cambridge: Cambridge University Press.

Wells, J. (2006). *English Intonation: An Introduction*. Cambridge: Cambridge University Press.

Wennerstrom, A. (1994). Intonational Meaning in English Discourse: A Study of Nonnative Speakers. *Applied Linguistics*, 15, 399–420.

Wennerstrom, A. (1998). Intonation as Cohesion in Academic Discourse: A Study of Chinese Speakers of English. *Studies in Second Language Acquisition,* 20, 1–25.

Xu, L. (2004). Manifestation of Information Focus. *Lingua, 114* (3), 277–299.

Xu, Y. (1999). Effects of Tone and Focus on the Formation and Alignment of F0 Contours. *Journal of Phonetics,* 27, 55–105.

Zielinski, B. W. (2008). The Listener: No Longer the Silent Partner in Reduced Intelligibility. *System,* 36, 69–84.

Received: 21 July 2022
Accepted for publication: 10 November 2022