The Influence of Gender on Prosodic Entrainment in Mandarin Conversations

XIA Zhihua, MA Qiuwu

Abstract The goal of this study is to find out how gender affects prosodic entrainment in Mandarin conversation. Based on the analyses of Tongji Games Corpus, it is found that in Mandarin conversations, mixed gender groups entrain on the greatest number of features and males entrain on the least; the entrainment degree is the most prevalent in mixed-gender groups, and the least in male groups. A cross-linguistic comparison between Mandarin Chinese and English finds striking similarities over the number of prosodic features and the degree of prosodic entrainment. The similarities support not only the view that entrainment is a cross-cultural phenomenon, but provide evidence that gender plays the similar role in prosodic entrainment in different language groups.

Key words Gender, Prosodic entrainment.

1. INTRODUCTION

The Conversations are considered as joint activities in which two interlocutors share or synchronize their mental states and performances. Prosodic entrainment means two speakers adapt prosody to that of their interlocutors in conversation, and then become similar in speaking for smooth and successful communication, which is also named prosodic accommodation, prosodic adaptation, or prosodic alignment.

According to Speech Accommodation Theory [6], the benefits of adapting communication to accommodate others are to increase communication efficiency and gain social approval or shorten desired level of social distance. Prosodic entrainment is important in social interaction. It assists the smooth expression and comprehension, and reveals the alignment of cognitive, expressive, and comprehensive layers in interaction, by which communication is fulfilled accurately and effectively ([3] [11] [12] [13]).

Prosodic accommodation varies according to the speakers' social status. For example, the participants having lower social status adapt more to the interlocutors having higher status [6]. Pardo [10] has examined the degree to which interlocutors increased the similarity in phonetic performances during conversational interaction, and find that both the role of a participant in the task and the sex of the pair of the talkers affected the degree of convergence. Levitan *et al*. [9] have accomplished research on the relationship between prosodic entrainment and social behavior, and find that mixed gender pairs entrain more than same gender pairs, and entrainment is more important to the perception of mixed gender pairs than it is for the same gender pairs.

It is difficult to control several social factors in one experiment simultaneously, although social factors cover various aspects. This paper focuses on the element of gender. The goal of this study is to find out how gender affects prosodic entrainment in Mandarin conversation.

In Section 2, the corpus and its annotation of this study is described. In Section 3, analyses of gender and prosodic entrainment are made. In Section 4, a cross-linguistic comparison of the gender and prosodic entrainment between Mandarin Chinese and English is made. In Section 5, conclusion, discussion, and the future research are provided.

2. CORPUS AND ANNOTATION

The analyses of this research are based on Tongji Games Corpus, which contains approximately 12 hours of spontaneous speech, each of which is about 6 minutes, task-oriented Mandarin conversations elicited by two games. Subjects are randomly selected from the students in Jiangsu Normal University with a National Mandarin Test Certificate level 2, with a grade of A or above. The requirements in the subjects' proficiency are to increase the likelihood that the Mandarin spoken in the corpus is standard, and reduce the influence of Chinese dialects in prosody.

70 pairs participated in the experiments, in which there are 23 female and female pairs, 30 female and male pairs, and 17 male and male pairs. A series of analyses over prosodic entrainment and gender are based on the conversations produced by these gender groups.

Because of the absence of some potential subjects, some of the available subjects participated twice in the experiments. In spite of this adjustment, the two people of any pair in conversation are still strangers. Therefore, 39 female subjects participated in the female-female conversations; 20 male subjects participated in the male-male conversations; 40 subjects (20 female subjects and 20 male subjects) participated in the mixed-gender conversations.

The recording of the corpus was accomplished in a sound-proof booth in Jiangsu Normal University. During the experiments, the two interlocutors faced two computers, and played the games presented on the screens. There was a curtain between them, so neither could see the other's facial expressions or body movements, in order to reduce the facial and gestures' aid in conversations, thus the function of prosody became prominent. Every speaker wore a head microphone (Sennheiser, PC166), and their conversations were recorded by another computer with Cool Edit (Pro. 2.0). The parameters in Cool Edit were set as 44100 HZ, 16, single track. Subjects accomplished free conversations, and nobody interrupted them in the course of the games till the end of the conversations. Subjects were told to play games only. They did not know the research purpose.

IPUs (Inter-Pausal Unit) are adopted as the minimal units in analyses. Casper [4] set the threshold of pause in IPU as 100ms in her research. Levitan & Hirschberg [8], Levitan *et al*. (9) set the threshold as 50ms. The threshold for IPUs of the present research is 80 ms by calculation. IPUs are automatically labeled by SPPAS [1]. And then, the IPUs' boundaries are checked manually in Praat.

The present research excluded the laugh, cough, sneezing, etc, which contained no linguistic contents from IPUs, and they are not annotated in Praat. The filled pauses, repairing, restarting, backchannel, etc, which contained linguistic contents, were included in IPUs, and they were considered as valid speaking and were annotated in Praat. Chinese characters are put down within IPUs in Praat instead of syllables.

The present research adopted the methods of Caspers [4] for identification of turns in Mandarin conversations.

The present research focuses on gender and prosodic entrainment in Mandarin conversations. Seven variables are set in the analyses of Tongji Games Corpus in this study. These parameters come from 3 main aspect of prosody, including the feature of duration (Speaking-rate), the features of F0 (F0 min, F0 mean, F0 max), and the features of intensity (Intensity min, Intensity mean and Intensity max).

Seven variables of every IPU were extracted by a Praat script. Data extraction is accomplished over the smallest analysis units---IPUs. However, some analyses of entrainment in the present research cover units larger than IPUs including turns, or conversations, which contain more than one IPU. The weighted averages are calculated over all the IPUs within these units.

3. THE ANALYSES OF GENDER AND PROSODIC ENTRAINMENT

The analyses of entrainment and gender include two parts: the number of prosodic features in entrainment and the entrainment degree in different gender combination. In the present research, there are three kinds of gender groups in conversation (female-female groups, female-male groups, male-male groups).

3.1 The number of prosodic features in entrainment

This analysis is to find out: Are there any differences in the number of prosodic features among different gender groups? The analyses in this section are set at the conversation level. Relevant studies mentioned in Section 1 have proved the influence of social factors in prosodic entrainment. Therefore, it is hypothesized that there should be differences in the number of prosodic features among different gender groups in Mandarin conversations.

3.1.1 Paired T-tests over different gender groups

Paired T-tests are accomplished between the partner distances and non-partner distances in this analysis. The partner distance is the distance of a prosodic feature between the speaker and his partner; non-partner distance is the mean of the distances of a prosodic feature between the speaker and other speakers, with whom he is not partnered in any conversations ([8] [9]). Non-partners in these games are restricted to those of the same gender and conversational role as their partners in dialogues. Thus, hypothesis of these analyses is that the partner distance should be smaller than the non-partner distance, which can supply the evidence for entrainment at the conversation level.

This method is explained by following formulas.

For each conversation, the present research defines *disp* as the partner distance between two partners (speaker A, speaker B) on the prosodic feature f :

$$disp = | A_f - B_f | \quad (1)$$

In Formula 1, *disp* represents the partner distance, *Af* and Bf are weighted average (as what mentioned in section 3) for the feature f over the whole conversation of the two partnered speakers A and B.

The present research define disnp as the non-partner distance on the feature f :

$$disnp = \frac{\sum_i | Af - Xif |}{| X |} \quad (2)$$

In Formula 2, *disnp* represents the non-partner distance, X(i) are the set of speakers, which are selected randomly in the Tongji Games Corpus. These speakers have the same gender and role as the speaker's partner, and are not paired with the speaker in any conversations. The restriction to the speakers with the same gender and role as the speakers' partner is to decrease the influence of gender and role in the results. *Af* and *Xif* are also the weighted mean (mentioned in Section 3) for the feature f over the whole conversation of the two non-paired speakers A and X. $| Af - Xif |$ represents the distance between non partners.

According to Section 2, in Tongji Games Corpus, the paired T-tests are accomplished

over the pairs with three kinds of gender combination (23female-female groups, 30 female-male groups, 17male-male groups). 39 female subjects participated in the female-female conversations; 20 male subjects participated in the male-male conversations; 40 subjects (20 female subjects and 20 male subjects) participated in the mixed-gender conversations.

For female-female groups, 39 pairs of partner distances and non-partner distances are calculated for one variable (one prosodic feature), and are put in a paired T-test. There are 7 variables in the present research, so 7 series of paired T-tests are accomplished over all the prosodic features examined.

For male-male groups, 20 pairs of partner distances and non-partner distances are calculated for one variable (one prosodic feature), and are put in a paired T-test. There are 7 variables in the present research, so 7 series of paired T-tests are accomplished over all the prosodic features examined.

For female-male groups, 40 pairs of partner distances and non-partner distances are calculated for one variable (one prosodic feature), and are put in a paired T-test. There are 7 variables in the present research, so 7 series of paired T-tests are accomplished over all the prosodic features examined.

3.1.2 Results

Since the paired T-tests are accomplished over 3 gender groups respectively, the results of analyses are listed separately.

For female-female group, the results of paired T-tests are listed in Table 1.

Table 1: Paired T-tests of female-female groups.

Feature	t	df	p-value	Sig.
Speaking rate	−5.792	38	0.0	*
F0 min	−1.283	38	0.207	/
F0 mean	−0.523	38	0.604	/
F0 max	−0.241	38	0.811	/
Intensity min	−1.726	38	0.092	/
Intensity mean	−4.765	38	0.0	*
Intensity max	−5.075	38	0.0	*

In Table 1, an asterisk * indicates the significant difference, and the symbol / indicates no significant difference. This table shows that in female-female conversations, speakers show significant entrainment over 3 prosodic features: Speaking-rate ($p= 0.0 < 0.05$), Intensity min ($p= 0.0 < 0.05$), and Intensity max ($p= 0.0 < 0.05$).

For male-male group, the results of paired T-tests are listed in Table 2.

In Table 2, an asterisk * indicates the significant difference, and the symbol / indicates no significant difference. Table 2 shows that in male-male conversations, speakers show significant entrainment over one prosodic feature: Speaking-rate ($p=0.0 < 0.05$).

Table 2: Paired T-tests of male groups.

Feature	t	df	p-value	Sig.
Speaking rate	−6.804	19	0.0	*
F0 min	0.453	19	0.656	/
F0 mean	0.228	19	0.822	/
F0 max	0.595	19	0.559	/
Intensity min	−1.869	19	0.077	/
Intensity mean	0.176	19	0.862	/
Intensity max	0.099	19	0.922	/

Table 3: Paired T-tests of female-male groups.

Feature	t	df	p-value	Sig.
Speaking rate	−3.024	39	0.004	*
F0 min	1.032	39	0.309	/
F0 mean	0.972	39	0.337	/
F0 max	−4.281	39	0.0	*
Intensity min	0.684	39	0.498	/
Intensity mean	−3.379	39	0.002	*
Intensity max	−3.225	39	0.003	*

For female-male group, the results of paired T-tests are listed in Table 3. In Table 3, an asterisk * indicates the significant difference, and the symbol / indicates no significant difference. Table 3 shows that in female-male conversations, speakers show significant entrainment over 4 prosodic features: Speaking-rate ($p=0.004 < 0.05$), F0 max ($p=0.0 < 0.05$), Intensity mean ($p=0.002< 0.05$)), and Intensity max ($p=0.003 < 0.05$).

The results of paired T-tests over the pairs with 3 kinds of gender combination show that in female-female conversations, speakers show significant entrainment over 3 prosodic features: Speaking-rate, Intensity min, and Intensity max; in male-male conversations, speakers show significant entrainment over 1 prosodic feature: Speaking-rate; in female-male conversations, speakers show significant entrainment over 4 prosodic features: Speaking-rate, F0 max, Intensity mean, and Intensity max.

Then it is found that more features are entrained in female-male conversations. The number of prosodic features entrained in mix gender group's conversations is the most, and the number of the prosodic features in male-male group's conversations is the least.

In conversations of the pairs with the same gender (female-female and male-male conversations), the features of duration and intensity are entrained; in mixed gender group's conversations, including the feature of duration, features of F0 are also entrained.

3.2 Entrainment degree of the pairs with different gender combination

From the results above in section 3.1, over Speaking-rate, Intensity mean and Intensity max, pairs in female-female, female-male and male-male conversations exhibit entrainment. It is necessary to explore further what is the difference of entrainment degrees over three gender combinations for anyone feature of the three.

The research question in this section is: Is there any difference in entrainment degree in the pairs of different gender combination? In addition to the number of features speakers entrained over, the present research also analyzes the degree of entrainment in the pairs with different gender combination.

3.2.1 ANOVA tests

ANOVA test is adopted in the analyses of entrainment degree, in which adaptation degrees are considered as dependent variables, and 3 gender combinations are independent variables.

The first step is to define a parameter—*sim*, in this analysis, which is calculated by Formula 3.

In Formula 3, *sim* represents the adaptation degree of one prosodic feature. *disp* represents the partner distance as in Formula 1, and *disnp* represents the non partner distance as in Formula 2 in Section 3.1.

$$sim = 1 - \frac{disp}{disnp} \quad (3)$$

In this analysis, *sim* represents the adaptation degree of the prosodic feature f. *disp* represents the partner distance as in Formula 1, and *disnp* represents the non partner distance as in Formula 2. In this analysis, non partner distance is considered as the baseline in the distance of the two interlocutors, the distance without the mutual adaptation. Thus, *disp/disnp*, *disp* normalized by *disnp*, represents the ratio of the remained distance after adaption between interlocutors to the primitive distance. Then, *sim*, *1-disp/ disnonp*, represents a similarity ratio, the percentage of becoming similar. This similarity ratio or normalization in Formula 3 has an advantage to control for speaker' differences.

For Speaking-rate, Intensity mean, and Intensity max, over which all the pairs with 3 gender combinations exhibit entrainment, ANOVA test is conducted over female-female, female-male, and male-male groups.

39 female subjects participated in the female-female conversations; 20 male subjects participated in the male-male conversations; 40 subjects (20 female, 20 male) participated in the mixed-gender conversations. *Sim* is calculated by Formula 3 over the values of *disp* and *disnp* from these conversations.

39 *sim* come from 39 female-female conversations; 20 *sim* come from 20 male-male conversations; 40 *sim* come from female-male conversations. ANOVA tests are accomplished over 3 prosodic features, taking *sim* as the dependent variable and 3 gender groups the independent variables.

3.2.2 Results

For Speaking-rate, Intensity mean, and Intensity max, over which all pairs with 3 gender combinations exhibit entrainment, ANOVA tests are conducted over the female-female, female-male, and male-male groups, and the results of these tests are listed respectively below.

The results of ANOVA test in terms of Speaking-rate over 3 gender combinations are listed in Table 4.

Table 4: An ANOVA test in terms of Speaking-rate over 3 gender combinations.

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between G-roups	2.06	2	1.028	4.508	.013
Within Groups	21.89	96	.228		
Total	23.94	98			

Table 4 shows that over Speaking-rate, *sim* has significant difference in the pairs with 3 gender combinations ($p=0.013 < 0.05$, $F= 4.508$). That is, in terms of Speaking-rate, entrainment degree is significantly different over the pairs with three gender combinations. It is necessary to test further the difference between every two gender combinations. Therefore, the multiple comparisons are made below.

The results of post-doc comparison are listed in Table 5.

Table 5 shows that over Speaking-rate, *sim* is the smallest in female-male conversations within 3 gender combinations, because *sim* in the female-male group is significantly smaller than that in the female-female group ($p=0.030< 0.05$) and *sim* in the female-male group is significant smaller than that in the male-male group ($p=0.007<0.05$). *Sim* is not significantly different between the female-female group and the male-male group. That is to say, in terms of Speaking-rate, the entrainment degree over the female-male group is the smallest, and there is no signify-cant difference between female-male group and the male-male group.

Table 5: Post-hoc comparison of ANOVA test in terms of Speaking-rate.

Multiple Comparison						
LSD						
(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
ff	fm	.23 622*	.10 745	.030	.0 229	.4 495
	mm	−.12 486	.13 132	.344	−.3 855	.1 358
fm	ff	−.23 622*	.10 745	.030	−.4 495	−.0 229
	mm	−.36 108*	.13 076	.007	−.6 206	−.1 015
mm	ff	.12 486	.13 132	.344	−.1 358	.3 855
	fm	.36 108*	.13 076	.007	.1 015	.6 206

The results of ANOVA test in terms of Intensity Mean in the pairs with 3 gender combinations are listed in Table 6.

Table 6: An ANOVA test in terms of Intensity Mean over 3 gender combination.

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.204	2	1.102	3.13	.048
Within Groups	33.791	96	.352		
Total	35.996	98			

Table 6 shows that over Intensity mean, sim is significantly different in the pairs with 3 gender combinations (p=0.048<0.05, F=3.131). That is, in terms of Intensity min, entrainment degree is significantly different over the pairs with three gender combinations. It is necessary to test further the difference between every two gender combinations. Therefore, the multiple comparisons are made below.

The results of post-hoc comparison are listed in Table 7.

Table 7: Post-hoc comparison of ANOVA test in terms of Speaking-rate.

Multiple Comparison						
LSD						
(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
ff	fm	.12 049	.13 351	.369	-.1 445	.3 855
	mm	.40 764*	.16 317	.014	.0 837	.7 315
fm	ff	−.12 049	.13 351	.369	−.3 855	.4 115
	mm	.28 715	.16 248	.080	−.0 354	.6 097
mm	ff	−.40 764*	.16 317	.014	−.7 315	−.0 837
	fm	.2 871	.16 248	.080	−.6 097	.0 354

Table 7 shows that over Intensity mean, sim is the smallest in male-male conversations within 3 gender groups, because sim in the male-male group is significantly smaller than in the female-female group (p= 0.014 < 0.05), and *sim* in the male-male group tends to be smaller than in the female-male group (p= 0.080 < 0.01). *Sim* in the male-male group is not significantly different from that in female-female groups or female-male groups. That is to say, in terms of Intensity mean, the entrainment degree over male-male group tends to be the smallest, and there is no significant difference between the female-female group and the female-male group.

The results of ANOVA test in terms of Intensity max in the pairs with 3 gender combinations are listed in Table 8.

Table 8: An ANOVA test in terms of Intensity Max over 3 gender combinations.

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.446	2	1.223	3.73	.028
Within Groups	31.484	96	.328		
Total	33.930	98			

Table 8 shows that over Intensity max, *sim* is significantly different in the pairs with 3 gender combinations ($p=0.028 < 0.05$, $F=3.729$). That is, in terms of Intensity max, entrainment degree is significantly different over the pairs with three gender combinations.

It is necessary to test further the difference between every two gender combinations. Therefore, the multiple comparisons are made below. The results of post-hoc comparisons are listed in Table 9.

Table 9 shows that over Intensity max, sim tends to be the smallest in male-male conversations among the 3 gender groups, because sim in the male-male group is significantly smaller than that in the female-female group ($p= 0.008 <0.05$), and sim in the male-male group is smaller than that in the female-male group, although the result is not significant ($p= 0.124 > 0.05$). Sim in the male-male group is not significantly different from that in the female-female group or the female-male group. That is to say, in terms of Intensity max, the entrainment degree in the male-male group tends to be the smallest, and there is no significant difference between the female-female group and female-male group.

According to the results of a series of ANOVA tests over Speaking-rate, Intensity mean, and Intensity max, it is found that in terms of Speaking-rate, the entrainment degree over female-male group is the smallest, and there is no significant difference over the female-male group and the male-male group; in terms of Intensity mean, the entrainment degree of the male-male group tends to be the smallest, and there is no significant difference between the female-female group and the female-male group; in terms of Intensity max, the entrainment degree of male-male group tends to be the smallest, and there is no significant difference between the female-female group and the female-male group.

Table 9: Post-hoc comparison of ANOVA test in terms of Intensity Max.

Multiple Comparison						
LSD						
(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
ff	fm	.18 379	.12 887	.157	−.0 720	.4 396
	mm	.42 689*	.15 750	.008	.1 143	.7 395
fm	ff	−.18 379	.12 887	.157	−.4 396	.0 720
	mm	.24 310	.15 683	.124	−.0 682	.5 544
mm	ff	−.42 689*	.15 750	.008	−.7 395	−.1 143
	fm	−.24 310	.15 683	.124	−.5 544	.0 682

Therefore, based on the results above, it is found that entrainment of mixed gender pairs is most prevalent although not necessarily strongest, male-male pairs tend to entrain least in MC conversations.

4. COMPARISON

A cross-linguistic comparison between Mandarin Chinese and English is made in this section. The comparison is mainly based on the results between the present research and that of Levitan *et al.* [9]. In terms of entrainment and gender, the cross-linguistic comparison covers

two aspects: prosodic features entrained by gender groups, and entrainment degree by gender groups.

4.1 Comparison of the prosodic features

The results of prosodic proximity over three gender groups are compared between Mandarin Chinese (MC) conversations in Tongji Games Corpus and Standard American English (SAE) in Columbia Games Corpus. The results of MC come from the analyses in Section 5.1, and the results of SAE come from the research of Levitan *et al*. [9]. The results of comparison are showed in Table 10.

In Table 10, the symbol "✓" represents showing similarity, "×" showing difference, and "—" showing no test on this feature.

Table 10 shows that the striking similarities are found in both languages over 3 gender groups. MC and SAE conversations show the similar patterns of entrainment over 3 gender groups. In detail, the mixed gender pairs entrain over the most of prosodic features, and male pairs over the least in both languages.

Table 10: The comparison of proximity over gender groups between MC and SAE.

	Female-Female		Male-Male		Female-Male	
	MC	SAE	MC	SAE	MC	SAE
Speaking-rate	✓	✓	✓	✓	✓	✓
F0 min	×	-	×	-	×	-
F0 mean	×	×	×	×	✓	✓
F0 max	×	×	×	×	✓	✓
Intensity min	×	-	×	-	×	-
Intensity mean	✓	✓	×	✓	✓	✓
Intensity max	✓	✓	×	✓	✓	✓

Only one difference is found in this analysis. For SAE, the prosodic features showing entrainment are consistent over 3 gender groups. That is, all the gender groups entrained over Intensity mean, Intensity max, and Speaking-rate.

But in MC, the prosodic features showing entrainment are not consistent over 3 gender groups, for example, male pairs entrain only on speaking rate.

4.2 Comparison of entrainment degree

The results of MC come from the analyses in Section 3, and the results of SAE come from the research of Levitan *et al*. [9].

According to the research of Levitan *et al*. [9], for SAE, it is found that entrainment on Intensity mean and max is the strongest for mixed gender pairs and the weakest for male pairs; the strength of entrainment on speaking-rate followed this pattern but the differences only approached significance ($p = 0.08$).

The similarity is found in MC and SAE conversations that the male pairs tend to entrain least.

The difference is found that in SAE conversations, entrainment is both strongest and most prevalent in mixed gender pairs, while in MC, it is the most prevalent in mixed gender pairs, but not necessarily the strongest.

5. CONCLUSION AND DISCUSSION

As a summary of the analyses, in the research on gender and prosodic entrainment, it is found that mixed gender pairs entrain over the greatest number of prosodic features, and male pairs on the least in both languages, that male-male pairs tend to entrain least, and that entrainment of mixed gender pairs is most prevalent although not necessarily strongest. Based on these results, conclusions can be made that in Mandarin conversations, in both

number of features entrained and degree of entrainment: males entrain the least.

These conclusions supported theories of dominance and perception partially. In relevant studies of entrainment and gender differences, the related theories are dominance and perception. In terms of dominance, Bilous & Krauss [2] make the male dominance hypothesis, and point out differences in speech between men and women have to do with men's dominant position in society. In addition, according to CAT [7], it is proposed that when there is an imbalance of power between speakers, the less dominant speaker will entrain more. In terms of perception, perception behavior link [5] states that perceiving something makes people more likely to mimic it and females are known to be more sensitive to perceived differences in speech. Based on these theories, predictions are made: in mixed gender pairs' conversations, females should entrain more, and there should be more entrainment in female-female pairs' conversations than male-male pairs' conversations. The result that the most prosodic features are involved in mixed gender pairs' entrainment, and in both number of features entrained and degree of entrainment, males entrain the least support partially the theories of dominance and perception.

A cross-linguistic comparison between Mandarin Chinese and English in terms of gender and prosodic entrainment is made. In the comparison the similarity in pattern between these two languages are striking: mixed-gender pairs entrain on the greatest number of features, and male pairs on the least; the most consistent results are for intensity mean, intensity max, and speaking rate, although all gender groups entrained on these in English, and male pairs entrain only on speaking rate in Mandarin.

The similarity of our findings in the relationship of gender and prosodic entrainment between Mandarin and English supports not only the view that entrainment is a cross-cultural phenomenon, but provides evidence that gender plays the similar role in prosodic entrainment in different language groups.

The present research analyzes the gender in entrainment. Actually, the entrainment in conversation is closely related to the social factors: the status, age, gender, role, region, ethnicity, environment, etc. The various responses to individual's interaction with the environment are assessed with reference to the individual himself; this assessment produces the perception of affordances which exist as a mix of an individual's abilities, their background, and even the environmental features.

Besides gender, more social factors should be involved in the study of prosodic entrainment. In order to make some factors prominent, it is necessary to control for other factors. The difficulty in research is how to appropriately control some factors and make others prominent in experiments. It is usually not easy to control several social factors in one experiment. Efforts should be made to produce well-balanced experimental design. Much future work could be done in this direction.

6. ACKNOWLEDGES

This work is supported by the Humanities and Social Sciences Foundation of China Education Ministry (15YJC740105) and 2015 Doctoral Project in Jiangsu Normal University (15XWR008). We thank Prof. Julia Hirschberg, Prof. Daniel Hirst and Dr. Rivka Levitan for their instructions and suggestions.

7. REFERENCES

[1] Bigi, B. & Hirst, D. 2012. Speech phonetization alignment and syllabification (sppas): a tool for the automatic analysis of speech prosody. In *Proceedings of Speech Prosody* 2012 (pp. 19–22). Shanghai: Tongji University.

[2] Bilous, F. R. & Krauss, R. M. 1988. Dominance and accommodation in the conversational behaviors of same- and mixed-gender dyads. *Language & Communication*, 8: 183–194.

Boylan, P. 2004. Accommodation theory revisited. Retrieved 10 Sep., 2012 from http://www.docin.com/p-688557528.html

[3] Caspers, J. 2003. Local speech melody as a limiting factor in the turn-taking system in Dutch. *Journal of Phonetics*, 31: 251-276.

[4] Chartrand, T. L. & Bargh, J. A. 1999. The Chameleon Effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6): 893-910.

[5] Giles, H., Coupland, J., Coupland, N. 1991. Accommodation Theory: Communication, Context, and Consequence. In H. Giles, J. Coupland & N. Coupland (eds.), *Contexts of Accommodation*. Cambridge: Cambridge University Press.

[6] Giles, H., Mulac, A., Bradac, J., Johnson, P. 1987. *Speech accom modation theory: the first decade and beyond*. Beverly Hills, CA: Sage.

[7] Levitan R. & Hirschberg, J. 2011. Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. *In Proceedings of Interspeech* (pp.3081–3084), Florence, Italy.

[8] Levitan R., A. Gravano, L. Willson, S. Benus, Hirschberg, J., Nenkova, N. 2012. Acoustic-prosodic entrainment and social behavior. In *Proceedings of Conference of the North American, Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 11-19). Montréal, Canada.

[9] Pardo J. S. 2006. On phonetic convergence during conversational interaction. *Acoustical Society of America*, 119(4): 2382-2393.

[10] Parrill, F. & Kimbara, I. 2006. Seeing and hearing double: the influence of mimicry in speech and gesture on observers. *Journal of Nonverbal Behavior*, 30(4): 157-166.

[11] Pickering M. J. & Garrod, S. 2004. Towards a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27: 169-226.

[12] Pickering M. J. & Garrod, S. 2006. Alignment as the basis for successful communication. *Research on Language and Computation*, 4: 203-228.

XIA Zhihua PhD She is a lecturer of Jiangsu Normal University. Her research interests include speech prosody, Intonation and L2 language acquisition.

E-mail: xzhlf@163.com

MA Qiuwu PhD He is a professor of Tongji University. His research interests include phonology, speech prosody, and phonetics.

E-mail: tjnkmqw@126.com